



**UNIVERSIDAD DE CHILE  
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS  
DEPARTAMENTO DE INGENIERÍA QUÍMICA Y BIOTECNOLOGÍA**

**MODELAMIENTO ESTRUCTURAL Y CARACTERIZACIÓN DE UNA LIPASA ACTIVA  
A BAJAS TEMPERATURAS MEDIANTE INGENIERÍA DE PROTEÍNAS**

**TESIS PARA OPTAR AL GRADO DE DOCTORA EN CIENCIAS DE LA  
INGENIERÍA, MENCIÓN QUÍMICA**

**FRANCESCA MERCADO MALEBRÁN**

**PROFESORES GUIA:  
JUAN A. ASENJO DE LEUZE  
BARBARA ANDREWS FARROW**

**MIEMBROS DE LA COMISION:  
ORIANA SALAZAR AGUIRRE  
INÉS CONTRERAS OSORIO  
ALVARO OLIVERA NAPPA**

**SANTIAGO DE CHILE  
2014**

## THESIS SUMMARY

Novel industries have incorporated the use of psychrophilic enzymes within their processes of production; this has allowed a reduction of substrate consumption and production of toxic compounds, energy use, costs and environmental impact, therefore increasing in the process yields. These enzymes have a high potential for the biotechnology industry, and among them, lipases are very important due to their high capacity to synthesize and to hydrolyze esters bonds and their high selectivity for substrates. Their applications involve food production, detergents, drugs, cosmetics, synthesis of chemical compounds such as biopolymers and agro-chemicals, and their use as biosensors, among many others.

In this work, a model of a cold-active lipase was determined using molecular modeling. This model provided a positive assessment which allowed characterizing this enzyme. Three amino acids in the catalytic triad: Ser 239, His 391 and Asp 361 were identified and their participation was studied, which presented a proper spatial distribution. To study more deeply the catalytic triad, a molecular *docking* method was used to predict coupling between the native lipase and 4 different substrates. The enzyme showed preference for 6 carbon acyl-chains substrates. In addition, molecular *docking* was used to determine the interaction between five different mutation of the catalytic triad and all substrates. To examine these mutations, models of the three dimensional structure of every single mutation were made and charge distributions and energies were analyzed. In all cases, interaction energies were dramatically lower in comparison with the native lipase, demonstrating the importance of this specific triad.

Residues studied were replaced by site-directed mutagenesis, and the effects on activity were assessed. In consequence all mutations completely abolished enzyme activity, proving ser239-asp361-his391 are the catalytic triad.

Finally, the phenomenon of interfacial activation accompanying the catalytic activity in a wide group of lipases and esterases was studied. Comparison with another enzyme model which has a similar secondary structure that acts as a lid, revealed two possible structures  $\alpha 8$  and G2 in the lipase studied that could be involved in this phenomenon.

To conclude, in this work the 3-D structure of this lipase was predicted and aligned with other lipases, which have known the 3-D structure to create an acceptable model. In addition, this work correctly predicted the amino acids belonging to the catalytic triad, Ser239, Asp361, and His391 that was also confirmed by site-directed mutagenesis. None of the mutants showed activity with the substrates tested. And finally, 3-D structure analysis showed two possible structures which can carry out the interfacial activation in the lipase studied.

## RESUMEN

Las nuevas industrias han incorporado el uso de enzimas psicrófilas dentro de sus procesos de producción, lo que les ha permitido reducir tanto el consumo de sustratos como la producción de compuestos tóxicos, además de ahorrar energía, permitiendo una disminución de costos e impacto ambiental y por ello un aumento en la rentabilidad del proceso. Estas enzimas tienen un alto potencial a nivel industrial biotecnológico, y dentro de éstas, las lipasas son foco de atención dada su alta capacidad de sintetizar e hidrolizar enlaces ésteres y su exquisita selectividad de sustratos. Sus aplicaciones involucran la producción de alimentos, detergentes, fármacos, cosméticos, síntesis de compuestos químicos como biopolímeros y agroquímicos y su uso como biosensores, entre otras.

En este trabajo, se determinó el modelo tridimensional de una lipasa activa a bajas temperaturas mediante modelamiento por homología. Este modelo presentó una evaluación positiva lo que permitió trabajar con él para caracterizar a esta enzima. Con este modelo se logró identificar y estudiar la participación de tres aminoácidos en la triada catalítica: Ser 239, His 391 y Asp 361, los cuales presentaban una acertada distribución espacial. Mediante *docking* molecular se estudió la unión y acoplamiento de estos tres aminoácidos con 4 tipos de sustratos diferentes, presentando preferencia por sustratos con cadenas aciladas de 6 carbonos. Se generaron mutaciones para cada uno de los aminoácidos candidatos, se realizaron modelos de su estructura tridimensional y se analizaron las distribuciones de cargas y energías. Así mismo, las mutaciones fueron sometidas a *docking* molecular para conocer la interacción de la triada mutada con estos cuatro sustratos. En todos los casos, las energías de interacción fueron drásticamente menores a la energía presentada por la lipasa nativa, demostrando la importancia de esta triada específica. Para validar estos resultados, cada uno de los residuos fue reemplazado utilizando mutagénesis sitio dirigida y se estudio su efecto en la actividad enzimática. Todas las mutaciones presentaron pérdidas de la actividad, probando que ser239-asp361-his391 corresponde a la triada catalítica. Finalmente, se estudió el fenómeno de activación interfacial que acompaña a la actividad catalítica en gran parte del grupo de lipasas y estererasas. La comparación con modelos de enzimas que presentaron estructuras secundarias tipo tapas, reveló dos posibles estructuras  $\alpha$ 8 y G2 en la lipasa estudiada, que podrían estar participando en este fenómeno.

Se concluye entonces, que la estructura de esta lipasa que fue predicha en base al alineamiento con otras lipasas de estructuras tridimensionales conocidas, corresponde al tipo  $\alpha/\beta$  hidrolasa. Además, este trabajo logró predecir correctamente que los aminoácidos que forman la triada catalítica son Ser239, Asp361, e His391 lo cual fue validado mediante mutagénesis sitio-dirigida, reflejando la falta de actividad de mutantes con la triada modificada. Finalmente, los análisis del modelo tridimensional creado mostraron dos posibles estructuras que podrían mediar la activación interfacial.

## TABLA DE CONTENIDO

<b>CAPÍTULO 1</b> .....	<b>10</b>
<b>1.1 INTRODUCCIÓN</b> .....	<b>11</b>
1.1.1 Biología de Sistemas.....	11
1.1.2 Ingeniería de Proteínas .....	12
1.1.3 Enzimas Industriales.....	12
1.1.4 Lipasas y sus aplicaciones industriales .....	13
1.1.5 Estructura y mecanismo de acción de las lipasas.....	14
1.1.6 Condiciones extremas y enzimas adaptadas a bajas temperaturas .....	18
1.1.7 Aplicaciones de enzimas adaptadas al frío.....	19
1.1.8 Adaptaciones enzimáticas a bajas temperaturas.....	19
<b>1.2 OBJETIVOS</b> .....	<b>20</b>
1.2.1 Objetivo General .....	20
1.2.2 Objetivos Específicos.....	20
<b>CAPÍTULO 2</b> .....	<b>22</b>
<b>2.1 RESUMEN</b> .....	<b>23</b>
<b>2.2 INTRODUCCIÓN</b> .....	<b>24</b>
2.2.1 Diseño Racional .....	24
2.2.2 Modelamiento de Estructuras de Proteínas .....	25
2.2.2.1 Modelamiento por Homología.....	25
2.2.2.2 Etapas en el modelamiento de estructuras de proteínas .....	26
<u>Identificación y selección del molde y del alineamiento</u> .....	26
<u>Construcción del alineamiento</u> .....	27
<u>Construcción del modelo</u> .....	27
<u>Evaluación del modelo</u> .....	28
<b>2.3 MATERIALES Y MÉTODOS</b> .....	<b>31</b>
2.3.1 Obtención de la estructura tridimensional de la lipasa 2-17 .....	31
2.3.1.1 Selección del molde.....	31
2.3.1.2 Alineamiento molde y obtención del modelo.....	31
2.3.2 Validación y refinamiento .....	31
2.3.2.1 Validación final del modelo tridimensional.....	32
<b>2.4 RESULTADOS Y DISCUSIÓN</b> .....	<b>33</b>
2.4.1 Búsqueda del molde.....	33
2.4.2 Alineamiento de la lipasa 2-17 con el molde 1jjiA .....	38
2.4.3 Construcción del modelo .....	40
2.4.4 Evaluación del modelo utilizando Verify 3D.....	41
2.4.5 Evaluación preliminar de modelo mediante el servidor VADAR.....	42
2.4.6 Refinamiento con Modeller .....	44
2.4.6.1 Selección de los aminoácidos para refinamiento.....	44
2.4.6.2 Refinamiento aminoácidos 259-288 .....	45
2.4.7 Evaluación mediante VADAR del modelo generado y refinado por Modeller .....	49
2.4.8 Comparación de los Gráficos de Ramachandran para el modelo tridimensional de la lipasa 2-17 antes y después del refinamiento. ....	52
2.4.9 Validación del modelo tridimensional de la Lipasa 2-17 utilizando el servidor MolProbity.....	54
2.4.10 Modelamiento a través de Swiss Model .....	55

2.4.11	Modelamiento a través de I-TASSER.....	57
2.4.11.1	Evaluación y selección de los modelos obtenidos por I-TASSER.....	59
2.4.12	Modelo final de la lipasa 2-17 .....	60
<b>2.5</b>	<b>CONCLUSIONES .....</b>	<b>65</b>
<b>CAPÍTULO 3 .....</b>		<b>66</b>
<b>3.1</b>	<b>RESUMEN.....</b>	<b>67</b>
<b>3.2</b>	<b>INTRODUCCIÓN .....</b>	<b>68</b>
3.2.1	Triada catalítica en lipasas .....	68
3.2.2	<i>Docking</i> Molecular .....	70
<b>3.3</b>	<b>MATERIALES Y MÉTODOS .....</b>	<b>71</b>
3.3.1	Determinación de la Triada Catalítica .....	71
3.3.1.1	Alineamiento de la secuencia de la Lipasa 2-17 con el molde.....	71
3.3.1.2	Análisis del modelo tridimensional de la Lipasa 2-17 .....	71
3.3.1.3	Alineamiento estructural del modelo tridimensional de la Lipasa 2-17 con su molde.....	71
3.3.2	Mutaciones de la Triada Catalítica .....	71
3.3.2.1	Modelamiento de las mutaciones de la triada catalítica de la Lipasa 2-17. ....	71
3.3.3	<i>Docking</i> Molecular .....	72
3.3.3.1	Evaluación de la unión del sustrato a la lipasa 2-17.....	72
3.3.3.2	Preparación de ligandos.....	72
3.3.3.3	Selección y ubicación de la caja .....	72
3.3.3.4	Generación de los mapas de búsqueda (Autodock) .....	72
3.3.3.5	<i>Docking</i> molecular.....	73
3.3.3.6	Evaluación de la unión dominio-sustrato de la lipasa 2-17 mutada .....	73
<b>3.4</b>	<b>RESULTADOS Y DISCUSIÓN.....</b>	<b>74</b>
3.4.1	Análisis de la estructura central de la lipasa 2-17.....	74
3.4.1.1	Plegamiento de la lipasa 2-17 .....	74
3.4.1.2	Aminoácidos involucrados en el mecanismo catalítico de la lipasa 2-17 .....	75
3.4.1.3	Comparación de las estructuras de lipasa 2-17 con la termófila AFEST .....	82
	Descripción de la triada catalítica de la carboxilesterasa 1JJI .....	82
	Comparación del sitio catalítico de la lipasa 2-17 con el de la carboxilesterasa 1JJI .....	83
3.4.1.4	Comparación de las estructuras de lipasa 2-17 con la psicrófila 2ORY.....	85
	Descripción de la triada catalítica de la psicrófila 2ORY.....	85
	Comparación del sitio catalítico de la lipasa 2-17 y la lipasa 2ORY.....	86
3.4.2	Estudio de la interacción lipasa nativa-triacilglicerol mediante <i>Docking</i> Molecular .....	88
3.4.3	Determinación de variantes para la triada catalítica de la lipasa 2-17 .....	93
3.4.3.1	Generación de Mutantes.....	93
3.4.3.2	Análisis computacional de las mutaciones.....	93
3.4.1	Estudio de la interacción de mutantes lipasa-triacilglicerol mediante <i>Docking</i> Molecular.....	98
3.4.2	Estudio de la interacción de mutantes lipasa-triacilglicerol de cadenas largas, mediante <i>Docking</i> Molecular.....	102
<b>3.5</b>	<b>CONCLUSIONES .....</b>	<b>107</b>
<b>CAPÍTULO 4 .....</b>		<b>108</b>
<b>4.1</b>	<b>RESUMEN.....</b>	<b>109</b>
<b>4.2</b>	<b>INTRODUCCIÓN .....</b>	<b>110</b>
4.2.1	Mutagénesis Racional .....	110
4.2.2	Mutagénesis Sitio Dirigida.....	111
4.2.3	Catálisis Enzimática .....	112

4.2.3.1	Flexibilidad Enzimática .....	112
4.2.3.2	Estabilidad Enzimática.....	114
<b>4.3</b>	<b>MATERIALES Y MÉTODOS .....</b>	<b>115</b>
4.3.1	Materiales .....	115
4.3.1.1	Cepas bacterianas .....	115
4.3.1.2	Reactivos .....	115
4.3.1.3	Vectores de clonamiento y expresión.....	116
	Vector de clonamiento pGEM-T Easy:.....	116
	Vector de expresión pMAL-c2E:.....	116
4.3.1.4	Medios de cultivo líquidos .....	117
	Medio Luria-Bertani (LB).....	117
	Medio TB.....	117
4.3.1.5	Medios de cultivo sólidos.....	118
4.3.1.6	Soluciones Stock.....	118
4.3.2	Metodología experimental.....	118
4.3.2.1	Cultivo de microorganismos.....	118
	Crecimiento y mantención de microorganismos en medio sólido .....	118
	Crecimiento de cepas en cultivos líquidos.....	118
4.3.2.2	Amplificación de ADN a través de PCR.....	118
4.3.2.3	Electroforesis de ADN.....	120
4.3.2.4	Purificación de ADN a partir de geles de agarosa .....	120
4.3.2.5	Ligación de fragmentos de ADN a vectores de clonamiento y expresión.....	120
	Ligación en el vector de clonamiento pGEM-T Easy.....	120
	Ligación en el vector de expresión pMAL-c2E .....	120
4.3.2.6	Transformación de células electrocompetentes.....	121
4.3.2.7	PCR de colonias .....	121
4.3.2.8	Minipreparación de ADN plasmidial .....	122
4.3.2.9	Digestión de ADN plasmidial .....	122
4.3.2.10	Inducción de la síntesis de proteínas recombinantes mediante el sistema de expresión <i>E. coli</i> /pMALc-2E.....	122
4.3.2.11	Purificación de la proteína recombinante mediante el sistema de expresión <i>E. coli</i> BL21 (DE3)/pMALc2E .....	123
4.3.2.12	Electroforesis en geles de poliacrilamida en condiciones denaturantes (SDS-PAGE).....	124
4.3.2.13	Medición de la actividad lipolítica de la proteína mediante ensayo en medio líquido .....	124
4.3.2.14	Determinación de la concentración de proteínas.....	124
<b>4.4</b>	<b>RESULTADOS Y DISCUSIÓN.....</b>	<b>125</b>
4.4.1	Mutación de la triada catalítica de la lipasa 2-17.....	125
4.4.1.1	Clonamiento de los mutantes para la triada catalítica de la lipasa 2-17 .....	125
4.4.1.2	Expresión de los mutantes para la triada catalítica de la lipasa 2-17 .....	129
4.4.1.3	Parámetros cinéticos de los mutantes de la triada catalítica.....	131
<b>4.5</b>	<b>CONCLUSIONES .....</b>	<b>135</b>
<b>CAPÍTULO 5 .....</b>	<b>137</b>	
<b>5.1</b>	<b>RESUMEN.....</b>	<b>138</b>
<b>5.2</b>	<b>INTRODUCCIÓN .....</b>	<b>139</b>
5.2.1	Activación interfacial.....	139
<b>5.3</b>	<b>MATERIALES Y MÉTODOS .....</b>	<b>142</b>
5.3.1	Análisis estructural de la lipasa 2-17 para determinar la activación interfacial .....	142

<b>5.4</b>	<b>RESULTADOS Y DISCUSIÓN.....</b>	<b>142</b>
5.4.1	Estudio de la activación interfacial de la Lipasa 2-17 .....	142
<b>5.5</b>	<b>CONCLUSIONES .....</b>	<b>148</b>
<b>CONCLUSIONES GENERALES .....</b>		<b>149</b>
<b>BIBLIOGRAFÍA.....</b>		<b>153</b>
<b>ANEXOS.....</b>		<b>162</b>
<b>8.1</b>	<b>ABREVIACIONES UTILIZADAS.....</b>	<b>163</b>
<b>8.2</b>	<b>MODELAMIENTO DE LA LIPASA 2-17 .....</b>	<b>163</b>
8.2.1	Secuencia lipasa 1jji.....	163
8.2.2	Alineamiento de las secuencias entregado por el archivo build_profile.ali: .....	164
8.2.3	Valores de corte entregados por el programa FUGUE.....	167
8.2.4	Evaluación del modelo utilizando Verify 3D.....	167
8.2.5	Refinamiento de los <i>loops</i> de la lipasa 2-17 .....	170
8.2.5.1	Archivo de entrada .....	170
8.2.5.2	Resultados Refinamiento .....	170
<b>8.3</b>	<b>IDENTIFICACIÓN DE LA TRIADA CATALÍTICA DE LA LIPASA 2-17 .....</b>	<b>174</b>
8.3.1	Secuencia lipasa 2-17 nativa .....	174
8.3.2	Secuencias mutaciones lipasa 2-17 .....	174
8.3.3	Energías de Afinidad Docking Molecular.....	175
<b>8.4</b>	<b>GENERACIÓN DE MUTANTES MEDIANTE MUTAGÉNESIS SITIO-DIRIGIDA.....</b>	<b>177</b>
8.4.1	Secuencia lipasa M37 .....	177
<b>8.5</b>	<b>ESTUDIO DEL FENÓMENO DE ACTIVACIÓN INTERFACIAL DE LA LIPASA 2-17 .....</b>	<b>177</b>
8.5.1	Curva de Progreso Lipasa 2-17 .....	177

## ÍNDICE DE TABLAS

Tabla I:	Tabla de valores de <i>z-score</i> de cada uno de los moldes.....	37
Tabla II:	Tabla resumen del análisis de la calidad de la estructura realizado con MolProbity.....	54
Tabla III:	Información del modelo obtenido por Swiss Model. ....	56
Tabla IV:	Estructuras secundarias de la lipasa 2-17 y sus aminoácidos correspondientes.....	63
Tabla V:	Cepas bacterianas utilizadas .....	115
Tabla VI:	Reactivos utilizados. ....	115
Tabla VII:	Partidores utilizados para la amplificación del gen de la lipasa 2-17 y para la amplificación de los mutantes.....	119
Tabla VIII:	Propiedades de cada una de las lipasas mutantes comparadas con la lipasa nativa. ....	132
Tabla IX:	Energía de activación <i>Docking</i> Molecular Lipasa nativa y diferentes sustratos.....	175
Tabla X:	Energía de activación <i>Docking</i> Molecular Lipasa nativa y variantes para sustrato de 6 carbonos. ....	175
Tabla XI:	Energía de activación <i>Docking</i> Molecular Lipasa nativa y variantes para sustrato de 8 carbonos. ....	176
Tabla XII:	Energía de activación <i>Docking</i> Molecular Lipasa nativa y variantes para sustrato de 10 carbonos.....	176
Tabla XIII:	Energía de activación <i>Docking</i> Molecular Lipasa nativa y variantes para sustrato de 10 carbonos.....	176

## ÍNDICE DE FIGURAS

Figura 1 : Reacción enzimática de una lipasa.....	13
Figura 2: Estructura tridimensional de una lipasa con un clásico plegamiento de $\alpha/\beta$ -hidrolasa.....	16
Figura 3: Diagrama del plegamiento $\alpha/\beta$ -hidrolasa.....	17
Figura 4: Gráfico de Lesk.....	26
Figura 5: Pasos en el modelamiento comparativo de estructuras de proteínas.....	29
Figura 6: Evaluación del modelo tridimensional de la lipasa 2-17 previo al proceso de refinamiento, realizada mediante la herramienta VADAR.....	43
Figura 7: Distribución de aminoácidos en la estructura secundaria de la lipasa 2-17.....	45
Figura 8: Valores molpdf entregados por Modeller al refinar el <i>loop</i> 259-288.....	46
Figura 9: Orientaciones para el <i>loop</i> formado desde el aminoácido 259 hasta el 288 una vez realizado el refinamiento.....	47
Figura 10: Modelo refinado de la lipasa 2-17 superpuesto con el modelo sin refinar de la lipasa 2-17.....	48
Figura 11: Comparación de modelos de la lipasa 2-17.....	49
Figura 12: Evaluación del modelo tridimensional de la lipasa 2-17 posterior al proceso de refinamiento, obtenida mediante la herramienta VADAR.....	51
Figura 13: Gráficos de Ramachandran del modelo tridimensional de la lipasa 2-17.....	53
Figura 14: Modelo tridimensional de la lipasa 2-17 obtenido mediante Swiss Model.....	57
Figura 15: Alineamiento de estructuras secundaria de los cinco modelos tridimensionales de la lipasa 2-17 obtenido por I-TASSER.....	59
Figura 16: Modelo tridimensional de la lipasa 2-17 obtenido mediante I-TASSER.....	60
Figura 17: Modelo tridimensional final de la lipasa 2-17 obtenido a través de Modeller.....	61
Figura 18: Plegamiento de la Lipasa 2-17. Arriba: Plegamiento característico de Lipasa 2-17.....	62
Figura 19: Estructura lipasa psicrófila M37.....	69
Figura 20: Alineamiento de secuencias entre la carboxilesterasa AFEST y la lipasa 2-17.....	75
Figura 21: Alineamiento estructural de la lipasa 2-17 y la cadena A de la lipasa AFEST.....	77
Figura 22: Modelo tridimensional de lipasa 2-17 con aminoácidos propuestos para el sitio activo.....	78
Figura 23: Plegamiento $\alpha/\beta$ hidrolasa.....	79
Figura 24: Esquema general del plegamiento $\alpha/\beta$ hidrolasa observado en la lipasa 2-17.....	80
Figura 25: Triada catalítica propuesta para la lipasa 2-17.....	80
Figura 26: Mecanismo catalítico de lipasas.....	81
Figura 27: Modelo y estereodiagrama del sitio catalítico de AFEST.....	82
Figura 28: Comparación de los aminoácidos del sitio activo de la carboxilesterasa 1JJJ y de la lipasa 2-17.....	84
Figura 29: Modelo y diagrama del sitio catalítico de la lipasa M37 (2ORY).....	86
Figura 30: Comparación de los aminoácidos sitio activo lipasa 2ORY y la lipasa 2-17.....	87
Figura 31: Esquema representativo de los triacilgliceroles usados para el <i>docking</i> molecular.....	88
Figura 32: Representación <i>Docking</i> Molecular.....	89
Figura 33: Interacción Triada catalítica lipasa 2-17 y triacilgliceroles.....	91
Figura 34: Triada catalítica de la lipasa 2-17 antes y después de la mutación S239N.....	94
Figura 35: Triada catalítica de la lipasa 2-17 antes y después de la mutación D361A.....	95
Figura 36: Triada catalítica de la lipasa 2-17 antes y después de la mutación D361E.....	96
Figura 37: Triada catalítica de la lipasa 2-17 antes y después de la mutación D361N.....	97
Figura 38: Triada catalítica de la lipasa 2-17 antes y después de la mutación H391N.....	97
Figura 39: Interacción mutante S239N y ácido caproico.....	99
Figura 40: Interacción mutante D361A y ácido caproico.....	100
Figura 41: Interacción mutante D361N y ácido caproico.....	100
Figura 42: Interacción mutante D361E y ácido caproico.....	101



Figura 43: Interacción mutante H391N y ácido caproico. ....	102
Figura 44: Acoplamiento de mutantes de la lipasa 2-17 y ácido caprílico.....	103
Figura 45: Acoplamiento de mutantes de la lipasa 2-17 y ácido cáprico.....	105
Figura 46: Acoplamiento de mutantes de la lipasa 2-17 y ácido palmítico .....	106
Figura 47: Esquema general Mutagénesis Sitio Dirigida. ....	111
Figura 48: Vector de clonamiento pGEM-T Easy.....	116
Figura 49: Vector de expresión pMAL-c2E.....	117
Figura 50: Mutagénesis sitio dirigida. ....	125
Figura 51: PCR 1 y 2 de la mutación S239N.....	126
Figura 52: PCR 1 y 2 de las mutaciones D361A, D361E y D361N. ....	127
Figura 53: PCR 1 fragmento directo H391N. ....	127
Figura 54: PCR final para cada mutación.....	128
Figura 55: Electroforesis de PCR de colonias. ....	129
Figura 56: Expresión del mutante S239N de la lipasa 2-17.....	130
Figura 57: Lipasa quimérica purificada para cada una de las mutaciones.....	130
Figura 58: Curva de progreso de la lipasa 2-17 y de los mutantes a temperatura a 20°C.....	131
Figura 59: Determinación de actividad a diferentes temperaturas.....	132
Figura 60: Determinación de actividad a diferentes pHs.....	133
Figura 61: Determinación de actividad catalítica de la lipasa sobre diferentes sustratos.....	134
Figura 62: Estructura cristalográfica de la lipasa PML. ....	140
Figura 63: Representación tapa $\alpha 3$ de la lipasa psicrofílica M37.....	143
Figura 64: Comparación de estructuras tridimensionales que participarían como tapas en Lip 2-17 y M37 (PDB ID 2ORY).....	143
Figura 65: Comparación de estructuras $\alpha$ hélices de la lipasa 2-17 vistas de manera lateral (izquierda) y frontal (derecha), que podrían estar actuando como tapa.. ....	144
Figura 66: Modelo de la lipasa 2-17 con $\alpha 1$ , $\alpha 8$ y G2.....	145
Figura 67: Esquema de la orientación de las estructuras que podrían actuar como tapa en la lipasa 2-17.. ....	146
Figura 68: Representación esquemática de la estructura primaria de la lipasa 2-17.....	147
Figura 69: Perfil de valores entregados por el programa Verify 3D.....	169
Figura 70: Curva de Progreso Lipasa 2-17 para el fenómeno de activación interfacial. ....	177

# **CAPÍTULO 1**

## **Introducción y Objetivos**

## 1.1 INTRODUCCIÓN

Las proteínas, ya sea directamente o indirectamente, son la base para la biotecnología moderna y en particular, el éxito en el diseño de proteínas que resulten ser útiles, está dado en gran medida por la reiterada experimentación. En un comienzo, los ingenieros de proteínas, con el objetivo de encontrar una proteína que presentara las características deseadas, se vieron obligados a elegir entre hacer librerías más grandes o bien enfocarse en librerías cuyas secuencias tuvieran resultados más efectivos, y hoy en día, uno de los principales objetivos es refinar esta capacidad para modificar una propiedad en particular. Tanto el diseño racional de proteínas, que emplea el conocimiento del mecanismo, la función y la estructura de las proteínas, así como los poderosos acercamientos del diseño no racional de proteínas, conocido como “Evolución Dirigida” han sido usados para generar proteínas con un alto éxito a nivel comercial, creando enzimas de gran potencial industrial (Cherry & Fidantsef, 2003). Así, estas enzimas han sido adaptadas para presentar mayor estabilidad, actividad, selectividad o incluso desarrollar nuevas funciones. Las nuevas tecnologías de ingeniería de proteínas continúan acelerando la tasa de descubrimiento y desarrollo de productos y enzimas que han comenzado a tener un gran impacto a nivel comercial (Acevedo y col., 2008; 2013). Esto posiblemente permitirá que una parte significativa de la economía global dependa de aplicaciones a partir de ingeniería de proteína tanto en agricultura, medicina e industrias alimenticias entre muchas otras, desarrollando procesos industriales verdaderamente sustentables.

### 1.1.1 Biología de Sistemas

La biología de sistemas (*System Biology*) utiliza un enfoque distinto basado en la modelización matemática de los procesos en estudio. Como resultado de la simulación, al poner a funcionar los modelos matemáticos con los que se representa al proceso, se obtiene una serie de predicciones del estado del proceso biológico que corresponderían a los resultados experimentales esperados. Durante las simulaciones, la red de interacciones entre los elementos que componen al proceso biológico, se representa con un sistema de ecuaciones diferenciales. Los valores de las características de esos elementos a distintos tiempos y bajo diversas condiciones experimentales (simuladas) son predecibles porque la dinámica, es decir los cambios del estado de ese sistema modelado, son calculables matemáticamente.

En este trabajo, los conceptos de Ingeniería de Proteínas y de Biología de Sistemas se han utilizado para analizar y mejorar una lipasa activa a bajas temperaturas, con el objetivo de aportar conocimiento y comprensión a este tipo de enzimas que hoy en día están siendo ampliamente utilizadas a nivel industrial.

### **1.1.2 Ingeniería de Proteínas**

El interés por modificar la funcionalidad de las proteínas, o incluso crear funciones nuevas, ha tomado cada vez más fuerza debido al gran número de aplicaciones industriales que requieren de enzimas como biocatalizadores, y que muchas veces deben presentar condiciones extremas, como eficiencia a muy bajas o muy altas temperaturas y pHs, tolerar agentes denaturantes u otros productos químicos, y trabajar en la ausencia de ligandos y proteínas accesorias.

Además, estas enzimas deben ser estables y estar adaptadas para las condiciones de trabajo industrial, sin dejar de lado su alta actividad específica y su especificidad por un sustrato dado. Asimismo, deben formar el producto deseado en los más altos niveles posibles, sin producir con ello compuestos tóxicos que pudiesen ser nocivos para la célula. Esto considerando que deben ser fáciles de producir y a bajo costo para que los procesos sean rentables.

Es así como la ingeniería de proteínas hoy en día ha permitido realizar un gran avance en la modificación de enzimas que tienen potencial industrial y, mediante los procesos de ingeniería racional de proteínas y evolución dirigida, muchas enzimas con potencial industrial han sido optimizadas.

Existen dos estrategias para diseñar proteínas que funcionen mejor bajo condiciones inusuales. La primera de ellas emplea mutagénesis sitio dirigida, que combinada con la información estructural de una proteína permite diseñar racionalmente una nueva y mejorada función. En contraste a esto, los acercamientos no racionales, conocidos como “Evolución Dirigida” utilizan técnicas de DNA recombinante para crear miles de posibles variantes que mediante métodos de selección de alto rendimiento, han permitido mejorar la búsqueda de nuevas proteínas (Giver y col., 1998). Este enfoque surgió como una poderosa alternativa para métodos racionales, particularmente cuando la relación entre la estructura y la función que se desea mejorar no es tan clara.

### **1.1.3 Enzimas Industriales**

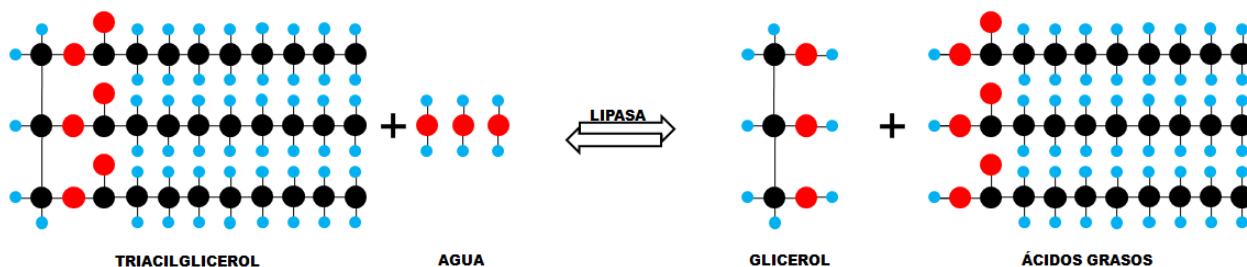
Las enzimas tienen un papel importante como catalizadores debido a su alto número de recambio y a sus refinados niveles de selectividad, y la utilización de éstas ha llevado a mejorar procesos y productos a nivel industrial. Hasta hace poco tiempo, la biodisponibilidad comercial de muchas enzimas era un factor que limitaba su uso, particularmente fuera de laboratorios especializados, pero hoy en día gracias a la expresión de genes foráneos en organismos como *Escherichia coli*, la manipulación de estos genes ha llegado a ser una técnica rutinaria, que ha asegurado que muchas más enzimas estén realmente disponibles y en cantidades suficientes para su estudio.

Por otra parte un gran número de genomas microbianos han sido secuenciados, lo que ha permitido conocer ciertos genes y con ello, conocer muchas enzimas de interés industrial, y junto con técnicas de mutagénesis que optimizan estas enzimas, ha crecido perceptiblemente la gama de enzimas que pueden ser usadas para aplicaciones industriales. Es así como enzimas nativas con baja especificidad a un sustrato, pobre estabilidad o insuficiente selectividad para una producción rentable, son actualmente modificadas y utilizadas para la producción en muchas áreas industriales como alimentos, empresas textiles, síntesis de químicos, biosensores y detergentes entre muchas otras.

#### 1.1.4 Lipasas y sus aplicaciones industriales

Las enzimas lipolíticas como lipasas, carboxilesterasas, y fosfolipasas participan en la degradación de diferentes tipos de lípidos. Las lipasas en particular catalizan la hidrólisis de triacilglicérols de cadena de largos variables entre 4 y 16 carbonos (Pleiss y col., 1998; Fojan y col., 2000). Todas estas enzimas son fundamentales para el funcionamiento de los organismos, tanto procariontes como eucariontes. Su relevancia se ve reflejada en las miles de publicaciones que aparecen cada año y en la versatilidad biotecnológica de estas enzimas dado su potencial para catalizar la degradación de lípidos. Dentro de las diferentes enzimas lipolíticas, las lipasas hoy en día representan una gran parte de las enzimas utilizadas a nivel industrial dadas sus múltiples aplicaciones en diversas áreas (Cavicchioli y col., 2002).

Una lipasa (triacilglicérol acilhidrolasas, E.C.3.1.1.3) es una enzima que bajo ciertas condiciones acuosas, es capaz de liberar ácidos grasos a partir de triacilglicérols, diacilglicérols y monoacilglicérols (Villeneuve y col., 2000; Jaeger Eggert, 2002). Su mecanismo de lipólisis consiste en la acción sobre los enlaces ésteres que se encuentran presentes en los acilglicérols liberando ácidos grasos y glicérol (Figura 1).



**Figura 1 : Reacción enzimática de una lipasa.** Hidrólisis de un triacilglicérol catalizada por una lipasa, resultando en la liberación ácidos grasos y glicérol como productos. Círculos celestes, rojos y negros representan hidrógenos, oxígenos y carbonos, respectivamente (Modificado de Villeneuve y col., 2000).

Los principales sustratos de las lipasas son triacilgliceroles de cadena larga que presentan baja solubilidad en agua, por ejemplo, acilgliceroles con cadenas de 10 ó más átomos de carbono (Jaeger y col., 1999; Villeneuve y col., 2000).

Las lipasas están ampliamente distribuidas en diversos animales, plantas y microorganismos, pero las lipasas microbianas son centro de atención debido a su enorme potencial en una gran variedad de aplicaciones biotecnológicas (Jaeger y col., 1999). Es por eso que microorganismos como bacterias, hongos y levaduras son las principales fuentes de estas enzimas.

Las lipasas se han clasificado de muchas maneras, siendo la más utilizada la clasificación de acuerdo a la secuencia. Las secuencias de lipasas microbianas y de hidrolasas homólogas son tomadas desde bases de datos; para la clasificación a nivel de superfamilias se incluyen todas las secuencias encontradas por iteraciones mediante la herramienta BLAST para una secuencia molde (Chothia & Lesk, 1986). Luego se subclasifican en familias mediante alineamientos de múltiple secuencias de las familias homólogas. Con ello, se logró establecer la existencia de 6 familias diferentes, entre las cuales, la familia IV, o grupo HSL (Hormone Sensitive Lipase, dada su gran similitud estructural con lipasas sensibles a hormonas de mamíferos), corresponde a lipasas adaptadas a bajas temperaturas. (Jaeger y col., 1999; Kanaya y col., 1998; Arpigny & Jaeger, 1999).

Para determinar que una enzima lipolítica corresponda a una lipasa verdadera se consideran dos factores: (i) la presencia de una cadena polipeptídica que actúa como tapa bloqueando el sitio activo, y (ii) por la activación e inactivación de la enzima dependiendo del solvente en el que se encuentra. A medida que el sustrato aciltriglicérido forma una emulsión, la enzima sufre un cambio conformacional y la tapa deja expuesto el sitio activo permitiendo el contacto con el sustrato. Cuando la lipasa no se encuentra en un solvente lipídico, esta cadena polipeptídica bloquea el sitio activo dejándolo inaccesible al sustrato (Jaeger y col., 1999; Brzozowski y col., 1990).

### **1.1.5 Estructura y mecanismo de acción de las lipasas**

Los triacilgliceroles son lípidos no cargados formados por una molécula de glicerol, que tiene esterificados sus tres grupos hidroxilo por tres ácidos grasos saturados o insaturados, y aunque los ácidos grasos de cadena corta son levemente solubles en agua, los compuestos con ácidos grasos de cadena larga esterificados a glicerol son insolubles (Gupta y col., 2004).

La actividad de una lipasa depende de la presencia de un solvente lipídico y es por ello que se definió una lipasa como “una carboxilesterasa que actúa en la emulsión de sustratos” (Desnuelle y col., 1960). Sarda y Desnuelle (Sarda & Desnuelle, 1958) demostraron cuantitativamente que la lipasa pancreática exhibe muy poca actividad cuando el sustrato, la triacetina (triacilglicerol constituido por

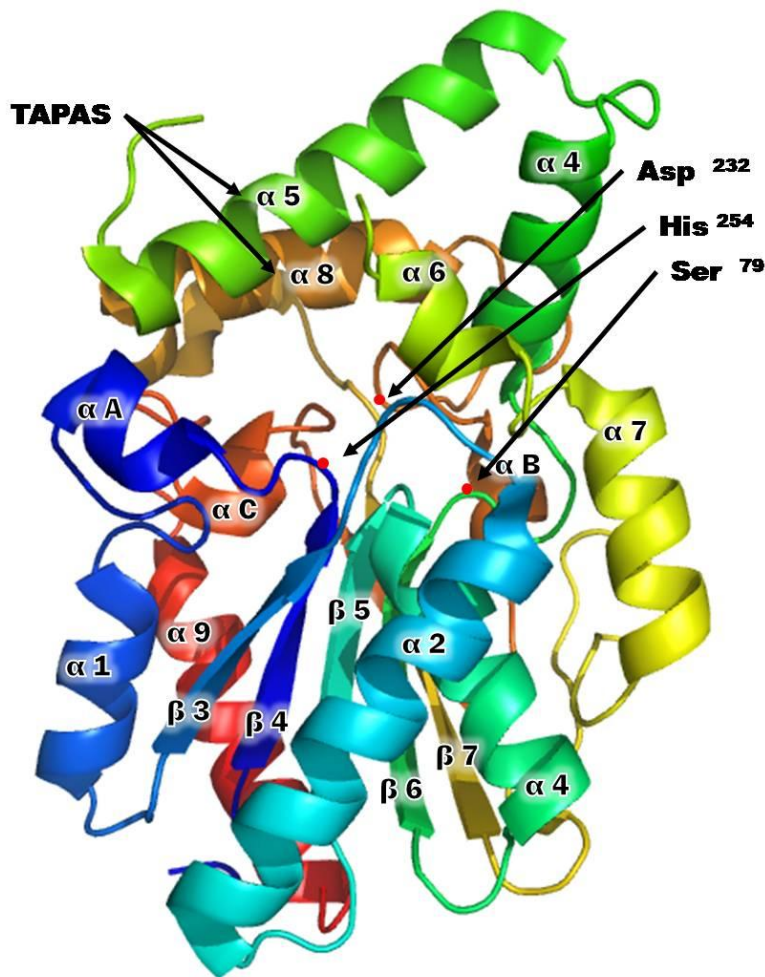
ácidos grasos de cadena corta) estaba perfectamente solubilizado. Sin embargo, cuando se utilizaban concentraciones de triacetina por encima de su límite de solubilidad, se formaban emulsiones que eran hidrolizadas por la lipasa de una forma más eficiente.

Este fenómeno de catálisis de una reacción heterogénea es único de las lipasas e involucra una activación interfacial, es decir, un incremento en la actividad enzimática cuando la solubilidad del sustrato se excede y aparece la fase lipídica. Esta propiedad trató ser explicada cuando fue dilucidada la primera estructura 3D de una lipasa. Se encontró que el sitio activo de la lipasa está cubierto por una cadena polipeptídica que actúa como tapa, la cual deja inaccesible el sitio activo para moléculas de sustrato (Winkler y col., 1990; Brady y col., 1990). Sin embargo, en la mayoría de los casos, cuando una lipasa se encuentra en la fase lipídica, un cambio conformacional toma lugar, causando que la tapa sea movida fuera y el sitio activo de la lipasa queda completamente accesible, permitiendo la hidrólisis del sustrato (Van Tilbeurgh y col., 1993; Sarda & Desnuelle, 1958). A comienzos de los años 90, gracias a la dilucidación de nuevas estructuras tridimensionales que permitieron obtener numerosos datos bioquímicos y cinéticos de las lipasas, se planteó la hipótesis de que no todas estas enzimas presentaban el fenómeno de activación interfacial, y que la presencia de la tapa polipeptídica no estaba necesariamente correlacionada con dicha activación (Sarda & Desnuelle, 1958). Hoy en día esta hipótesis sigue en estudio, pero aun no es claro cuál es el real mecanismo que permite la activación de las lipasas.

### **Estructura de las lipasas**

Las lipasas microbianas son extensamente utilizadas para catalizar reacciones de hidrólisis, alcoholisis, esterificación y transesterificación de triacilgliceroles y de otros ésteres insolubles en agua ya que presentan una amplia especificidad de sustrato.

A pesar de que muchas lipasas catalizan las mismas reacciones, al mirar sus secuencias es posible observar que éstas varían ampliamente, sin embargo, todas conservan un pentapéptido, el motivo GxSxG, donde el nucleófilo es siempre una serina. Este motivo forma un doblez tipo  $\gamma$  entre la lámina  $\beta_5$  y la  $\alpha$  hélice siguiente, lo que forma la estructura más conservada del llamado plegamiento  $\alpha/\beta$  hidrolasa (Figura 2), muy característico de las lipasas (Pleiss y col., 2000). La maquinaria catalítica de las enzimas con este tipo de plegamiento está formado por una triada catalítica, que corresponde a tres residuos ubicados bajo una geometría muy conservada en los *loops* y en el sitio de la hoja  $\beta$ . Así, un residuo nucleófilo que suele ser una serina, un residuo catalítico comúnmente aspartato o glutamato y finalmente un residuo de histidina forman la triada (Jaeger y col., 1999). La histidina catalítica se ubica en un *loop* (de largo y conformación variable) justo después de la hoja  $\beta_8$  del plegamiento  $\alpha/\beta$  hidrolasa (Fischer & Pleiss, 2003).



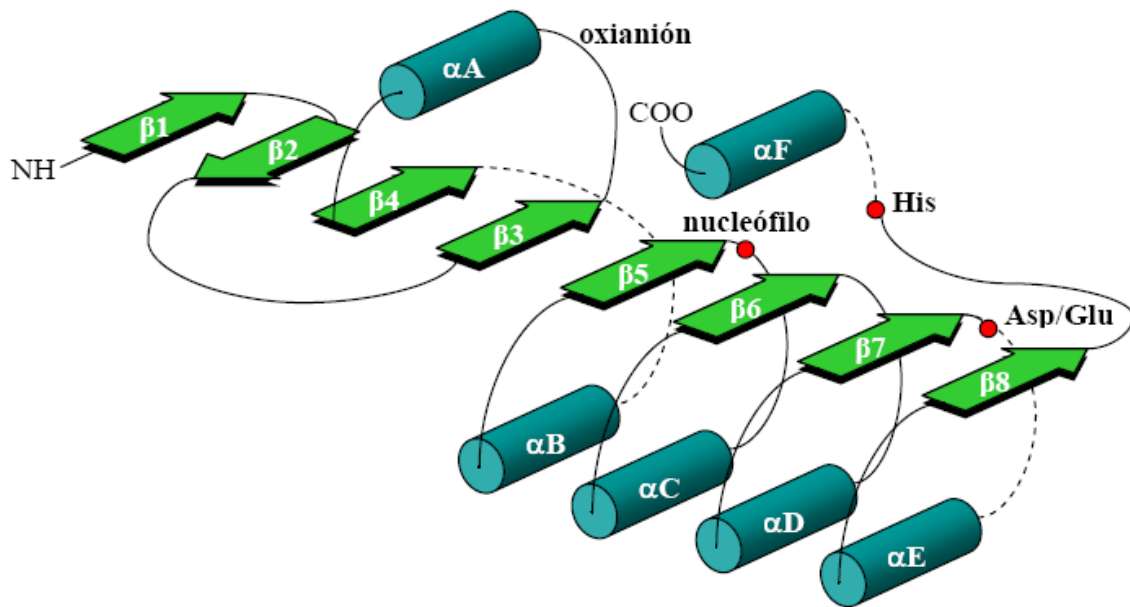
**Figura 2: Estructura tridimensional de una lipasa con un clásico plegamiento de  $\alpha/\beta$ -hidrolasa.** Las hélices  $\alpha$  y las hojas  $\beta$  fueron coloreadas de acuerdo a la posición en la secuencia, desde el N-terminal en azul hasta el C-terminal en rojo. La triada catalítica conformada por serina, aspartato e histidina se exhibe a través de las abreviaciones Ser79, Asp232 e His254 y se encuentran marcadas con un círculo rojo. Las hélices  $\alpha$  que actúan como tapas son indicadas en la figura. Figura realizada con el programa Pymol (Figura basada en Korman & Bowie, 2012) (PDB ID 4GW3).

La estructura  $\alpha/\beta$  hidrolasa está conformada por ocho hojas  $\beta$  centrales entre las cuales siete se encuentran ubicadas de forma paralela, mientras que solo una ( $\beta_2$ ), de manera antiparalela. Las hojas  $\beta$  conforman una estructura semicilíndrica (Figura 3).

Las estructuras  $\beta_1$  y  $\beta_2$  se encuentran unidas directamente y de manera continua. La hoja  $\beta_2$  se une con la  $\beta_3$  pero de manera interrumpida por la hoja  $\beta_4$ , que a su vez está conectada con la  $\beta_3$  mediante una  $\alpha$  hélice. La hoja  $\beta_4$  se conecta a través de una  $\alpha$  hélice con la  $\beta_5$  (Figura 3). Este patrón se repite hasta la última hoja  $\beta$ , y la estructura finaliza con una  $\alpha$  hélice que contiene el carboxilo terminal (Gupta y col., 2004).



Asociado al sitio activo se encuentra la cavidad oxianiónica que está formada por parte de la una cadena polipeptídica, específicamente un residuo situado en la región N-terminal y un residuo en la región C-terminal próximo al residuo nucleofílico catalítico (Jaeger y col., 1999; Pleiss y col., 2000). El residuo de la cavidad oxianiónica que se encuentra próximo al nucleófilo es estructuralmente conservado y puede formar parte de dos secuencias consenso (Pleiss y col., 2000). Es importante considerar que las lipasas se diferencian entre sí en la posición de las  $\alpha$  hélice como también de los *loops* que entran en contacto con el sitio activo (Nardini & Dijkstra, 1999).



**Figura 3: Diagrama del plegamiento  $\alpha/\beta$ -hidrolasa.** La hoja  $\beta$  está representada por flechas verdes mientras que las hélices  $\alpha$  lo están por cilindros de color azul. Las líneas discontinuas indican la localización de posibles inserciones. Los círculos rojos representan la ubicación de los residuos de serina, (nucleófilo) ubicado posterior a la hoja  $\beta_5$ , el residuo catalítico aspartato o glutamato situado generalmente a continuación de la hoja  $\beta_7$  y el residuo de histidina altamente conservado ubicado a continuación de la última hoja  $\beta$ . (adaptado de Jaeger y col., 1999).

Por otra parte, el mecanismo catalítico se ve favorecido por la presencia de bolsillos en el sitio activo, que acomodan las cadenas acilo de sus sustratos. Estas hendiduras son denominadas *sn-1*, *sn-2*, *sn-3*. Así, la preferencia por el sustrato está determinada por el carácter hidrofóbico del sitio activo y el tamaño de los bolsillos (Jaeger y col., 1999).

## **Catálisis y especificidad de la lipasa al sustrato**

El mecanismo de catálisis se inicia en un primer paso, cuando el sustrato interactúa con el residuo nucleofílico del sitio activo, formándose un intermediario tetraédrico transitorio, que es estabilizado por los residuos de la cavidad oxianiónica, la histidina y el residuo ácido del sitio activo. La transferencia de un protón por parte de la histidina produce la liberación del alcohol y la formación de un intermediario covalente. Luego, el ataque de un nucleófilo, como agua en una reacción de hidrólisis, alcohol o éster en reacciones de transesterificación, forma nuevamente un intermediario tetraédrico. Así, la transferencia de un protón lleva a la formación de producto, ácido o éster, y a la regeneración de la lipasa (Jaeger y col., 1999).

### **1.1.6 Condiciones extremas y enzimas adaptadas a bajas temperaturas**

La gran diversidad de condiciones climáticas en conjunto con diversos factores geográficos origina una variedad de ambientes, muchos de los cuales resultan inhóspitos e incluso extremos para los seres humanos y, sin embargo, se encuentran colonizados por un sinnúmero de organismos. La mayor parte de estos organismos son considerados extremófilos, es decir, organismos capaces de habitar ambientes extremos, y corresponden mayoritariamente a procariontes. Los microorganismos extremófilos que habitan lugares de altas temperaturas son llamados termófilos y para los que viven en zonas de bajas temperaturas criófilos o psicrófilos (Jaenicke & Závodszy, 1990; Jaenicke, 1991; Jaenicke, 1998; Hough y col., 1999).

La temperatura es uno de los factores ambientales más importantes para la vida. Dentro de las diferentes condiciones de temperatura, los ambientes fríos están ampliamente distribuidos (Russell y col., 1998; Cavicchioli y col., 2000). Se piensa que cerca de un 85% de la biosfera terrestre experimenta temperaturas menores a los 5°C, ocurriendo esto en continentes Ártico y Antártico, regiones montañosas y agua marina que cubre un 70% de la superficie (Gerday y col., 2000; Sellek & Chaudhuri, 1999).

Estos microorganismos que habitan ambientes extremos, deben llevar a cabo procesos celulares que han sido adaptados para su crecimiento y sobrevivencia a estas bajas temperaturas. Todos los componentes de la membrana celular y de sistemas de transporte de solutos intracelulares, ácidos nucleicos y proteínas deben ser adaptados convenientemente (Marhuenda-Egea & Bonete, 2002; Demirjian y col., 2001). Otros procesos celulares fundamentales como el metabolismo, replicación, transcripción y traducción de proteínas deben también estar diseñados para soportar el frío (Feller, 2003; 2007). Además, estos microorganismos han desarrollado la capacidad de impedir la formación de estructuras cristalinas de agua intracelular (Cavicchioli y col., 2002).

Las bacterias acostumbradas a vivir bajo estas condiciones cuentan con un sistema enzimático que trabaja en frío. Es así como una de las estrategias de tolerancia y adaptación al frío más ampliamente utilizada entre los diferentes organismos es la modificación de los parámetros cinéticos, principalmente de aquellas enzimas fundamentales en las vías metabólicas. Otro mecanismo de adaptación es la síntesis de nuevas enzimas o isoenzimas que se encuentran activas a bajas temperaturas, la síntesis de enzimas chaperonas que ayudan en el plegamiento de otras enzimas, una disminución en la tasa de degradación de proteínas que participan en las principales rutas metabólicas y en algunos casos un decrecimiento de la tasa metabólica (Arpigny y col., 1997; Feller & Gerday, 1997; Gianese y col., 2001).

Además, las enzimas que se han ido adaptando al frío presentan temperaturas a las que se obtiene el máximo de actividad desplazadas cerca de los 15°C y termolabilidad a temperaturas moderadas, entre 20°C y 60°C (Gerday y col., 1997; Feller & Gerday, 2003; Russel, 1998). Es así como muchos de los organismos adaptados a bajas temperaturas han desarrollado diversos mecanismos para sobrevivir bajo estas condiciones extremas.

#### **1.1.7 Aplicaciones de enzimas adaptadas al frío**

Las bacterias criofílicas están acostumbradas a vivir bajo condiciones de muy bajas temperaturas, por lo que cuentan con un sistema enzimático que trabaja en frío el cual es responsable de su adaptación. Así, los genes que codifican estas enzimas, son cada vez más estudiados y más aplicaciones están siendo atribuidas a ellas. Dentro de las principales aplicaciones industriales de las enzimas adaptadas al frío, se encuentran i) la industria textil con procesos como pulido biológico ii) la industria de detergentes en la cual la incorporación de estas enzimas criofílicas podría disminuir el consumo energético iii) la industria alimenticia, principalmente lácteos, jugos, carnes, cerveza y vinos iv) la industria química, en la cual estas enzimas median procesos que involucran reactivos y productos sensibles a temperaturas v) en biotecnología permitiendo procesos como biorremediación de lugares con condiciones extremas, la inactivación de productos fundamentales en áreas como biología molecular que utilicen otras enzimas, y la elaboración de biosensores vi) y finalmente en la industria farmacéutica para la preparación de productos de cosmetología y medicamentos (Cavicchioli y col., 2002).

#### **1.1.8 Adaptaciones enzimáticas a bajas temperaturas**

Muchos representantes de bacterias gram positivas y negativas, arqueas, levaduras, algas y microalgas han sido encontradas en ambientes de muy bajas temperaturas, y exhiben además una serie de adaptaciones a bajas temperaturas (Margesin, y col., 2002). Estos microorganismos adaptados a bajas

temperaturas se ven enfrentados a reto termodinámico de mantener las reacciones catalizadas por enzimas y sus tasas metabólicas con un crecimiento cerca o bajo de punto de congelamiento del agua pura (D'Amico y col., 2003; Wintrode & Arnold, 2000; Marx y col., 2007). Sin embargo, en ambientes naturalmente fríos, el consenso de la estrategia adaptativa de las proteínas criofílicas consiste básicamente en perder la estabilidad, obteniendo una proteína más móvil o flexible a temperaturas que “congelarían” los movimientos moleculares y las velocidades de reacción de proteínas mesófilas (Rothschild & Mancinelli, 2001; Van der Burg, 2003).

Las estructuras cristalográficas de este tipo de enzimas pueden exhibir adaptaciones moleculares para la temperatura, particularmente, el incremento del número y la fuerza de las interacciones débiles y de factores estructurales que están involucrados en la estabilidad de proteínas desde psicrófilos a mesófilo y a termófilos (Kumar y col., 2001; Lopez y col., 2007; Byun y col., 2007). De hecho, el carácter dominante de una enzima adaptada a bajas temperaturas es probablemente su número de recambio mejorado, es decir el número de reacciones elementales que realiza la enzima por cada centro activo en una cierta unidad de tiempo y su eficacia catalítica (Lonhienne y col., 2000), que compensan la reducción de la velocidad de reacción a bajas temperaturas con el fin de mantener los flujos metabólicos adecuados (Lonhienne y col., 2001; Russell y col., 1998; Siddiqui & Cavicchioli, 2006). Sin embargo, la mayoría de las adaptaciones celulares y mecanismos moleculares no son entendidos completamente y constituyen uno de los principales focos de estudios de adaptación (Feller, y col., 1997).

## **1.2 OBJETIVOS**

### **1.2.1 Objetivo General**

Identificar aspectos determinantes en la catálisis de una lipasa activa a bajas temperaturas mediante la obtención de un modelo útil de su estructura tridimensional, que permita conocer y caracterizar esta enzima mediante herramientas bioinformáticas e ingeniería de proteínas.

### **1.2.2 Objetivos Específicos**

- Diseñar y construir mediante modelamiento por homología, un modelo tridimensional de la lipasa 2-17 cuya evaluación resulte positiva.
- Identificar residuos potencialmente involucrados en la actividad lipolítica, particularmente los aminoácidos de la triada catalítica característica de este tipo de enzimas.

- Mediante *docking* molecular, estudiar la interacción de la triada propuesta con sustratos preferentes por enzimas lipolíticas.
- Generar mutantes para los aminoácidos propuestos como candidatos de la triada, y estudiar mediante *docking* molecular su acoplamiento con los sustratos analizados en el punto anterior.
- Validar los resultados obtenidos *in silico* mediante la construcción y caracterización bioquímica de los mutantes generados para los aminoácidos catalíticos, utilizando mutagénesis sitio-dirigida.
- Identificar posibles estructuras presentes en el fenómeno de activación interfacial y estudiar su relación con la catálisis enzimática de esta lipasa.

## **CAPÍTULO 2**

**Modelamiento de una Lipasa Activa a Bajas Temperaturas**

## 2.1 RESUMEN

La importancia de los métodos computacionales utilizados en el diseño racional radica en que éstos pueden ayudar al análisis y la comprensión de los resultados experimentales obtenidos en biología molecular. El objetivo de los modelos creados para representar las estructuras de las proteínas, es básicamente poder predecir cómo se disponen los átomos de una proteína en base a su secuencia conocida, con el fin de estudiar su función molecular, su historia evolutiva e incluso, si el modelo es correcto, diseñar o muestrear ligandos y calcular sus afinidades.

Particularmente, el modelamiento por homología es una herramienta que surge de la necesidad de conocer estructuras no determinadas de proteínas con secuencias conocidas. Este modelamiento consta de cuatro etapas. (i) La primera etapa es determinar homólogos de la proteína blanco. Posteriormente, (ii) se realiza un alineamiento entre la proteína blanco y los “homólogos”, para ver que regiones posiblemente tienen estructura común. (iii) Una vez realizado el alineamiento se procede a construir el modelo y (iv) finalmente los distintos modelos generados deben ser evaluados para ver si el modelo se ajusta a los datos experimentales y si tiene capacidad predictiva.

En este capítulo, se realizó el modelamiento por homología de la estructura tridimensional de la lipasa antártica 2-17. Se comenzó con la búsqueda de proteínas homólogas a la proteína blanco. Para ello, se envió la secuencia de aminoácidos de la enzima al programa FUGUE que seleccionó diez secuencias molde cuyas estructuras fueron determinadas por difracción de rayos X. A partir de las diez secuencias, se seleccionó como molde una carboxilesterasa, que presentó la mejor evaluación hecha por el programa Modeller. A partir de su archivo PDB se originaron cinco modelos posibles para la lipasa 2-17, dentro de los cuales el mejor evaluado fue el modelo 3. Este modelo fue escogido como estructura tridimensional de la lipasa 2-17, y fue nuevamente evaluado mediante herramientas como Verify 3D, Procheck, VADAR y MolProbity entre otros. Para analizar y validar el modelo obtenido por Modeller, se realizaron modelamientos utilizando los servidores Swiss Model e I-TASSER, los cuales entregaron modelos tridimensionales de la lipasa 2-17 con una estructura muy similar a la obtenida con Modeller, y finalmente un refinamiento de la estructura obtenida para mejorar el posicionamiento de los *loops*. En todos los casos, la estructura de la lipasa correspondió a una estructura tipo  $\alpha / \beta$  hidrolasa formada por una hoja  $\beta$  compuesta por 8 hebras y rodeada por hélices  $\alpha$ , formando una estructura tipo barril.

## 2.2 INTRODUCCIÓN

El principal objetivo del modelamiento por homología de la estructura de una proteína es construir un modelo tridimensional (3D) de una proteína cuya estructura es desconocida, en base a la comparación de la secuencia de una proteína de estructura conocida. Para construir un modelo útil, se deben cumplir dos condiciones: (i) Primero, la similitud de secuencia entre la proteína molde (proteína de estructura desconocida) y el molde (proteína de estructura conocida) deben ser considerable (sobre un 40% de identidad), y (ii) segundo, el alineamiento entre la secuencia de la proteína blanco y la secuencia de la proteína molde debe ser correctamente calculado.

El modelamiento comparativo es posible debido a que pequeños cambios en la secuencia usualmente resultan en pequeños cambios en la estructura 3D. Aunque se han hecho progresos considerables en la predicción de estructuras de proteínas, el modelamiento comparativo de estructuras de proteínas sigue siendo el método de predicción más exacta. La precisión global de los modelos comparativos se extiende ampliamente, así desde modelos de baja a modelos más precisos comparables a estructuras determinadas por cristalografía o resonancia magnética nuclear (RMN). Incluso modelos de baja resolución pueden ser útiles en biología debido a que algunos aspectos de la función a veces pueden ser predichos en base a las características estructurales secundarias de un modelo.

Particularmente, el diseño racional es una estrategia en ingeniería de proteínas que intenta crear proteínas mejoradas basadas en la estructura tridimensional y la relación entre la estructura y la función, la cual ha sido desarrollada por años como parte de la ciencia de proteínas.

### 2.2.1 Diseño Racional

El diseño racional fue una de las primeras aproximaciones al concepto de ingeniería de proteínas y sigue siendo ampliamente utilizado, ya sea como un enfoque independiente o bien, combinado con mutagénesis al azar y/o evolución dirigida.

Herramientas como cristalografía de rayos X y RMN que entregan la estructura tridimensional de una proteína, pueden ser usadas en base a procesos iterativos, para representar modelos computacionales. Este modelo computacional creado permite realizar predicciones, y particularmente es utilizado para determinar efectos de mutaciones en las propiedades basadas en la estructura de una proteína.

Una de las principales propiedades del diseño racional es la capacidad de crear variantes de una proteína nativa usando técnicas de ADN recombinante. El método de ADN recombinante más utilizado es mutagénesis sitio dirigida en la cual un aminoácido de una posición particular es sustituido por otro. Posteriormente se selecciona y se purifica la nueva proteína, y finalmente se realiza una evaluación de la



propiedad estudiada. La estructura de la nueva proteína es utilizada para explicar, a través de modelamiento molecular, los resultados de la elección inicial y proporciona un nuevo cambio que puede ser utilizado en una segunda ronda de mutagénesis.

En problemas donde la relación entre estructura y función es sencilla, el diseño racional es una muy buena opción. Pero las propiedades tales como la estabilidad e incluso algunas cuya relación con la estructura es aun menos obvia, tienen mejores resultados utilizando además algunos enfoques y técnicas de evolución dirigida (Duilio y col., 2004, Cherry & Fidantsef, 2003).

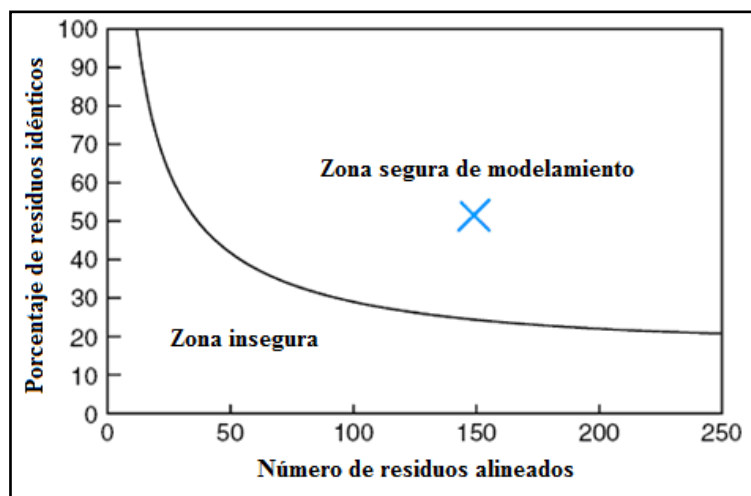
## **2.2.2 Modelamiento de Estructuras de Proteínas**

Los nuevos métodos de modelamiento pueden hoy en día dar comprensión detallada de reacciones catalizadas por enzimas, incluyendo un análisis de los mecanismos enzimáticos y la identificación de factores determinantes de la especificidad y la eficiencia catalítica de una enzima (Lüthy y col., 1992). Es así como la enzimología computacional ha emergido como un nuevo campo que contribuye significativamente con el diseño basado en estructuras de proteínas y el desarrollo de moldeos predictivos (Guex y col., 1999; Hinsen, 2000).

### **2.2.2.1 Modelamiento por Homología**

El modelamiento por homología es una metodología utilizada para construir modelos útiles de la estructura tridimensional de una proteína cuya estructura real es desconocida. Para construir el modelo se necesita conocer la secuencia aminoacídica de la proteína y la estructura tridimensional de alguna proteína homóloga, evolutivamente relacionada, obtenida por difracción de rayos X o por resonancia magnética nuclear (NMR).

Este método se basa principalmente, en dos observaciones: (i) la estructura de una proteína se encuentra únicamente determinada por su secuencia de aminoácidos, por lo tanto, conocer la secuencia de una proteína debería ser suficiente para obtener su estructura (Epstain y col., 1963), y (ii) durante la evolución, la estructura tridimensional es más conservada que la secuencia de aminoácidos asociada y cambia mucho más lentamente. Así, secuencias similares poseen estructuras prácticamente idénticas e incluso proteínas lejanamente relacionadas presentan plegamientos muy similares (Chothia & Lesk, 1986). Así, fue posible determinar un límite preciso en el cual el modelamiento era confiable. Este límite se estableció en función del número de aminoácidos alineados y el porcentaje de identidad de secuencia (Figura 4) (Rost, 1999).



**Figura 4: Gráfico de Lesk.** Si la relación entre el número de residuos y el porcentaje de identidad de dos secuencias se ubica dentro de la zona marcada como segura, es posible establecer que las dos secuencias poseerán un plegamiento muy similar (Modificado de Rost, 1999).

### 2.2.2.2 Etapas en el modelamiento de estructuras de proteínas

El modelamiento por homología se basa en cuatro etapas principales: (i) búsqueda y selección del molde, (ii) construcción del alineamiento entre la secuencia blanco y el molde, (iii) construcción del modelo que incluye la generación del esqueleto de la proteína, y el modelamiento de *loops* y cadenas laterales, y finalmente (iv) la evaluación del modelo.

#### Identificación y selección del molde y del alineamiento

La selección del mejor molde se encuentra íntimamente relacionada con la construcción del alineamiento entre las secuencias blanco y molde, ya que su identificación radica en la producción de alineamientos de secuencia. Éste es un paso trascendental dentro del modelamiento por homología, ya que la calidad del modelo a obtener depende en gran medida de la calidad del alineamiento seleccionado.

Existen diversas metodologías para la identificación de los posibles moldes, que pueden ser divididas en tres clases. La primera clase incluye los métodos que comparan la secuencia blanco con cada secuencia disponible en la base de datos independientemente, es decir, por pares y que sólo consideran información proveniente de la secuencia. A este tipo pertenecen programas como FASTA y BLAST. El segundo conjunto de métodos se basa en la comparación de múltiples secuencias, lo que aumenta la sensibilidad del método de búsqueda. Un ejemplo de este tipo de metodología es PSI-BLAST, en el cual se determina un primer conjunto de homólogos, se construyen los alineamientos múltiples y finalmente se obtiene una matriz de puntajes que se utiliza para la búsqueda de nuevos homólogos.

Esto se realiza iterativamente hasta que no se encuentran más homólogos (Sander & Schneider, 1991). Finalmente el tercer conjunto de métodos se denominan “método de encaje con moldes tridimensionales”, y se basan en la comparación entre la secuencia blanco y una secuencia de una proteína con estructura conocida. Esto se realiza de manera independiente para cada par de secuencias blanco-molde. Dentro de este tipo de métodos se encuentra el programa FUGUE que para predecir si la secuencia blanco adoptará o no la estructura tridimensional de alguna proteína, el programa realiza una optimización del alineamiento con respecto a una función de puntaje cuyo valor depende de la estructura (Shi y col., 2001).

La selección del molde más adecuado para la construcción de un modelo depende del objetivo para el cual la proteína está siendo modelada. Es por ello, que mientras mayor sea la similitud entre la proteína a modelar con la proteína molde, mejor será el molde. De la misma manera, mientras mayor sea la proximidad evolutiva entre ambas proteínas el modelo será más preciso (Chothia & Lesk, 1986). Por otra parte, se deben comparar los ambientes en los que la proteína molde y la proteína blanco adquieren su estructura cuaternaria, considerando factores como solventes, pH, ligandos y otras interacciones. Finalmente se debe considerar la calidad de la estructura tridimensional disponible para el molde, por ejemplo, analizando el factor R si la estructura de la proteína molde fue obtenida por cristalografía o el número de restricciones por residuo si la estructura fue obtenida por resonancia magnética nuclear (NMR) (Sander y col., 1991).

### Construcción del alineamiento

Una vez escogido el molde, se procede con la construcción del alineamiento. Para ello es importante considerar que si el porcentaje de identidad de secuencia entre la proteína blanco y el molde es superior al 40%, pueden aceptarse como válidos los alineamientos construidos, pero al disminuir el porcentaje de identidad, aumenta el número de gaps o porciones donde la secuencia blanco no se alinea con el molde. Para ello, existen programas que corrigen automáticamente los alineamientos, como el programa CLUSTAL (Martí-Renom y col., 2000).

### Construcción del modelo

Al igual que para la selección del molde, existen tres tipos de metodologías para la construcción de modelos a partir de un alineamiento. El método más ampliamente utilizado es “*Modeling by assembly of rigid bodies*”. Este método construye el modelo a partir de un pequeño número de cuerpos rígidos (zonas cuyas coordenadas relativas entre sí se mantienen fijas al ser sometido a fuerzas externas) obtenidos desde la estructura de la proteína alineada (Martí-Renom y col., 2000).

El segundo tipo de metodologías disponibles para la construcción de modelos se denomina “*Modeling by segment matching or coordinate reconstruction*” y se basa en que la mayoría de los segmentos formados por seis péptidos de la estructura de las proteínas pueden ser agrupados en aproximadamente 100 clases estructurales. Así, los modelos por homología se construyen ensamblando segmentos cortos que siguen como guía a un subconjunto de posiciones atómicas, obtenidas desde la estructura molde. Usualmente los carbonos alfa del esqueleto de la proteína son utilizados como guía (Martí-Renom y col., 2000). Finalmente la tercera clase de metodologías se denomina “*Modeling by satisfaction of spatial restraints*”. Estos métodos generan una serie de restricciones o limitaciones sobre la estructura que puede adoptar la secuencia blanco asumiendo que los ángulos y largos de enlace deben ser similares a los que poseen los residuos alineados con el molde. Las restricciones son de tipo estereoquímicas y se aplican sobre el largo de enlaces, ángulos de enlace, ángulos dihedrales y contactos entre átomos no unidos. El modelo se obtiene mediante la minimización de las infracciones a estas restricciones (Martí-Renom y col., 2000). El programa Modeller pertenece a este tipo de metodología.

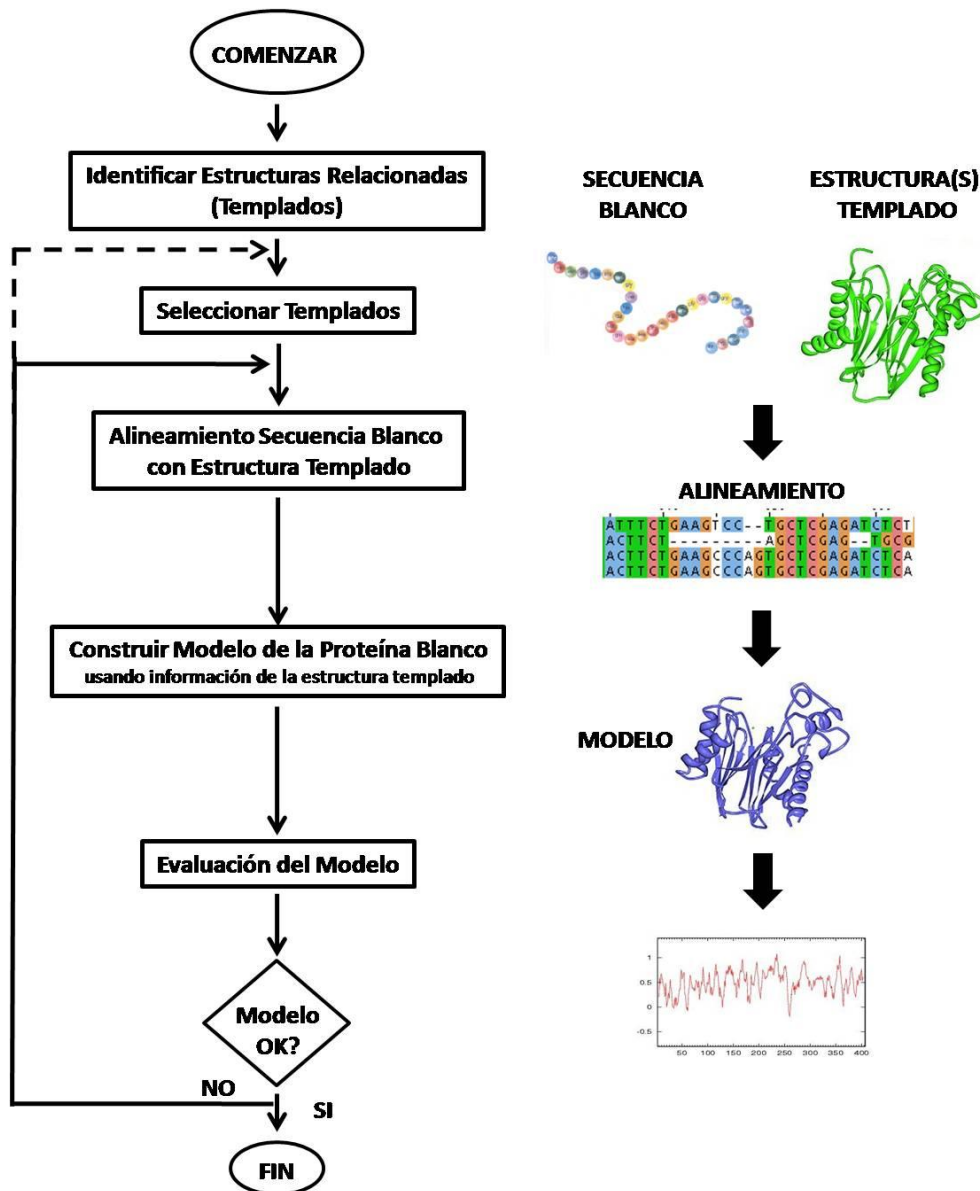
En general, estos métodos modelan primordialmente la estructura de la cadena principal de una proteína. Sin embargo, existen otras zonas que deben ser consideradas y que presentan mayores dificultades al momento de modelar, como son los *loops* que unen estructuras secundarias, y las cadenas laterales (Martí-Renom y col., 2000). Los *loops* usualmente participan en la función específica de una proteína, por lo que resultan de trascendental importancia cuando el modelo a construir va a ser utilizado en el estudio de interacciones ligando-proteína. La principal limitante del modelamiento de *loops* radica en que su conformación tridimensional debe obtenerse, en gran parte, sólo con la información obtenida desde su propia secuencia y dado que en general son muy cortos, ésta suele ser insuficiente para asegurar que la conformación obtenida es la real (Fiser y col., 2000; Gerstein & Krebs, 1998).

### Evaluación del modelo

A medida que la similitud entre la secuencia blanco y secuencia molde disminuye, los errores en el modelo aumentan. Los principales errores de un modelo son (i) errores en el empaquetamiento de las cadenas laterales, y (ii) distorsiones y cambios en regiones correctamente alineadas. Por ejemplo, la conformación de la cadena lateral puede cambiar como consecuencia de las diferencias entre las secuencias, aunque el plegamiento global siga siendo el mismo. Así, es posible que en regiones correctamente alineadas, el molde sea localmente diferente del blanco lo que resulta en errores en esta región. Algunas veces estas discrepancias no se deben a diferencias entre las secuencias, sino más bien, a errores en la utilización de las técnicas para la determinación de la estructura molde. Esto se puede minimizar utilizando múltiples moldes. (iii) Además se deben considerar errores en regiones donde no

hay alineamiento con un molde, por ejemplo por inserciones o deleciones, (iv) errores debido a malos alineamientos, que corresponde a la fuente más importante de errores, especialmente cuando la identidad de secuencia es menor que el 30%, y finalmente (v) errores por la utilización de un molde incorrecto.

Es por ello, que es necesaria la evaluación del modelo antes de extraer cualquier tipo de información y seguir cada una de las etapas para el modelamiento de proteínas (Figura 5). Para esta evaluación existen muchos servidores y programas disponibles.



**Figura 5: Pasos en el modelamiento comparativo de estructuras de proteínas.** Una vez identificados los moldes, se selecciona el más adecuado, se realiza el alineamiento con la estructura escogida y se construye el modelo. Finalmente el modelo debe ser evaluado (Modificado de Martí-Renom y col., 2000).

El primer paso en la evaluación es asegurarse que el modelo presente el plegamiento correcto. Su exactitud puede ser estimada en base a dos criterios: (i) una alta similitud con el molde más cercano, es decir mediante su potencial DOPE (*Discrete Optimized Protein Energy*) (Shen & Sali, 2006) o por la conservación de residuos claves en la secuencia blanco. Además, debe poseer una adecuada estereoquímica. Existen algunos programas útiles para evaluar la estereoquímica de un modelo tales como Procheck, Aqua, Whatcheck y Squid, siendo el más utilizado Procheck. Este tipo de programas analiza largos y ángulos de enlace, planaridad del enlace peptídico y de los anillos presentes en las cadenas laterales, quiralidad, ángulos de torsión de la cadena principal y de las cadenas laterales, entre otros.

Otros criterios que pueden considerarse para la evaluación de un modelo corresponden a la distribución de características espaciales como: formación del núcleo hidrofóbico, accesibilidad al solvente de residuos y átomos, distribución espacial de grupos cargados, distribución de las distancias entre átomos, volúmenes atómicos, y enlaces de hidrógeno en la cadena principal. Comparando los valores obtenidos del modelo para estas propiedades con los valores obtenidos para estructuras de proteínas determinadas con altas resoluciones, se puede tener una idea si el modelo es correcto o no, dado que no se esperan grandes desviaciones en estos valores para todas las proteínas (Martí-Renom y col., 2000).

## **2.3 MATERIALES Y MÉTODOS**

### **2.3.1 Obtención de la estructura tridimensional de la lipasa 2-17**

La estructura tridimensional de la lipasa en estudio (lipasa 2-17) se obtuvo mediante la realización de un modelo por comparación. Se utilizó para esto la secuencia aminoacídica de la proteína completa (Anexo 8.3.1). Todas las estructuras fueron visualizadas con el programa Pymol (DeLano, 2002).

#### **2.3.1.1 Selección del molde**

Se realizó la búsqueda de las secuencias conocidas de proteínas que presentaran un correcto alineamiento a la secuencia objetivo, con el fin de utilizarlas como moldes para elaborar el modelo por comparación. Esta búsqueda fue realizada mediante la herramienta FUGUE y en la base de datos del programa Modeller 9.14 (Eswar y col., 2007). Además la búsqueda fue complementada con la información entregada por la base de datos de Protein Data Bank (Bernstein y col., 1977).

#### **2.3.1.2 Alineamiento molde y obtención del modelo**

Mediante el programa Modeller se elaboró un modelo tridimensional de la lipasa 2-17, que se realizó en base al alineamiento entre la secuencia de la proteína blanco que quiere ser modelada y la proteína molde cuya estructura ha sido resuelta previamente (Sánchez & Sali, 2000).

El molde utilizado para el modelamiento comparativo fue la carboxilesterasa AFEST (PDB ID 1JJI) (Anexo 8.2.1), la cual fue alineada con la secuencia objetivo. Posteriormente se generaron 200 modelos a partir de este alineamiento. Utilizando el índice DOPE score como criterio de selección se escogió el modelo con menor índice.

### **2.3.2 Validación y refinamiento**

El modelo realizado con el software Modeller fue evaluado, posteriormente refinado y finalmente validado. Para el proceso de evaluación y validación del modelo previo al refinamiento, se utilizó la el servidor en línea VADAR (Willard y col., 2003), con el objetivo de detectar problemas asociados a los residuos de la proteína modelada. Para reparar los errores detectados se utilizó la herramienta “*loop refinement*” del programa Modeller.

Se generaron 100 refinamientos y se seleccionó el mejor modelo, como aquel que presentó menor DOPE score.

### **2.3.2.1 Validación final del modelo tridimensional**

La validación final del modelo se llevó a cabo graficando los valores DOPE obtenidos para el modelo antes y después de ser refinado y comparando ambos perfiles (datos no mostrados). Esta validación se realizó mediante el servidor VADAR que permitió la comparación de los problemas presentes en el modelo, antes y después de llevar a cabo el refinamiento. Además se realizó la gráfica de Ramachandram con el servidor MolProbity (Davis y col., 2004), para determinar si los residuos están ubicados en zonas con o sin impedimento estérico.



## 2.4 RESULTADOS Y DISCUSIÓN

### 2.4.1 Búsqueda del molde

La búsqueda de los moldes se realizó con el programa FUGUE. Este programa reconoce diferentes homólogos mediante la comparación entre la secuencia y la estructura. Para ello, utiliza tablas de sustituciones ambiente específicas y penalizaciones de gap dependientes de la estructura, donde los puntajes de cada apareamiento de aminoácidos y la relación inserción/delección son evaluadas dependiendo del ambiente local de cada residuo aminoacídico en una estructura conocida. Dada la secuencia molde, FUGUE busca los perfiles estructurales disponibles en las bases de datos, calcula los puntajes de compatibilidad estructura-secuencia y elabora una lista de potenciales homólogos y alineamientos.

El programa FUGUE entregó diez secuencias (de diez familias diferentes) de proteínas homólogas para la secuencia de la lipasa 2-17 (Anexo 8.2.2), con un nivel de confianza del 99% al utilizar el valor de corte recomendado por el programa (z-score=6.0) (Ver tabla de valores de corte en Anexo 8.2.3).

```
>P1;Query
sequence:Query:::::0.00: 0.00
MPILPVPSLNALLTKTLKTVKKTAAKNVHQHHILHHTLKGLDNLPTLLERVNHQLKASTAEQYPLADAH
LRLILAVSNKLRPLALDKLPNLRQKFGTDAVSLQAPSVWQQNAEASGENSENAVSWQDKTIANADGGDMT
VRCYQOSKQNNNEGKSTDEVAMLFFHGGGFCIGDIDTHHEFCHTVCAQGTGWAVVSVDYRMAPEYPAPTALK
DCLSAYAWLTHEQSLSGALPSRIVLSGDSAGGCLAALVAQQVIKPIDALWQDNNQATETDKKANDTFKKS
LADLPRPLAQPLYPVTDYEAEPYSEWELYGEG-LLLDHNDANVFNTAYTQHSGLMQSHPLISVMHGDNAH
LCPSYVVVAELDILRDEGLAYAEELLQKEGVKVVETTYTVLGAPHGFINLMSVHQGLGNQNTNYIIKSFACFVQ
NLLTSEGDEPNL*
```

Lo primero para comenzar a trabajar es escribir las secuencias entregadas por FUGUE en formato PIR para que pueda ser leído por Modeller. El archivo anterior se conoce con la extensión **.ali**. La primera línea contiene el código de la secuencia de la lipasa 2-17 en el formato "**>P1;code**" y corresponde a la línea de identificación. La segunda línea está compuesta por diez campos separados por dos puntos, y contiene información sobre los archivos de estructura.

El primer campo indica si existe o no una estructura tridimensional para la proteína cuya secuencia se presenta más abajo y el método utilizado para su obtención (structureX, Rayos X). El segundo campo corresponde al código asignado a la proteína en la base de datos de Protein Data Bank. El tercer y cuarto campo, identifican el número de residuo y la cadena a la que pertenece el primer residuo del alineamiento. El quinto y sexto campo indican el número de residuo y la cadena a la que pertenece el último aminoácido del alineamiento. El resto de los campos son optativos, y establecen el nombre de la proteína, el microorganismo fuente, la resolución cristalográfica con que fue obtenido el modelo, y el factor R del análisis cristalográfico, respectivamente. Los campos que no se identifican se dejan en blanco. Finalmente, cada secuencia debe terminar con un asterisco (\*).

La búsqueda de las posibles secuencias relacionadas a una estructura conocida puede ser realizada mediante el comando **profile.build()**. El programa Modeller entrega un archivo "**build\_profile.py**" el cual también es entregado en otros formatos. Como se muestra más adelante, el archivo **build\_profile.prf** contiene información sobre los parámetros de entrada utilizados por Modeller para construir los perfiles.

En este caso, las columnas más importantes en el archivo **build\_profile.prf** corresponden a la primera, la séptima y la octava. La primera columna entrega el código de la secuencia PDB que será comparada con la secuencia *target*. La séptima columna entrega el porcentaje de identidad de secuencia entre la lipasa y la secuencia PDB normalizada por el largo del alineamiento. En general, un valor de identidad de secuencia sobre aproximadamente 25% indica un potencial molde a menos que el alineamiento sea muy corto (menos de 100 residuos). Una mejor medida del significado del alineamiento es el *e-value*, dado en la octava columna. En este caso cuatro secuencias PDB muestran similitudes muy significantes a la secuencia de la lipasa con *e-value* iguales a 0. El *e-value* corresponde al "valor de expectación" e indica la probabilidad de que por azar se encuentre una secuencia determinada en las bases de datos. Así, un valor bajo de *e-value* para un alineamiento indica una baja probabilidad de haber encontrado por azar dicha secuencia (mientras menor es el *e-value* mayor significancia tiene el valor del *score*).

Para seleccionar el molde más apropiado para el modelamiento de la lipasa 2-17 a partir de estas cuatro estructuras, se utiliza el comando **alignment.compare\_structures()** que determina la similitud entre la secuencia y la estructura de los posibles moldes. Se elabora el archivo **compare.py** (mostrado abajo) para que Modeller realice la evaluación.

```

1HITS FOUND IN ITERATION:      1
Dynamically allocated memory at      amaxprofile [B,KiB,MiB]:      7637332      7458.332
7.284
> 1ea5A      1  329  4500  532  431  26.42      0.60E-02  2  173   65  255   18
210
> 2bce      1  507  5400  532  431  28.00      0.51E-04  3  138  110  255   51
200
> 1qe3A      1  1619 5450  467  431  27.69      0.33E-04  4  116  161  276   90
219
> 1u4nA      1  2934 16150 308  431  40.93      0.0       5  233  136  402   55
291
> 1f6wA      1  3121 5200  533  431  26.75      0.15E-03  6  139  110  255   51
207
> 1jjiA      1  5080 18800 311  431  34.84      0.0       7  286   88  410   18
304
> 1jkmA      1  5099 10650 358  431  31.71      0.0       8  275   86  392   41
327
> 1k4yA      1  5362 4550  501  431  21.79      0.42E-02  9  129  112  242   47
202
> 1lzlA      1  6028 13000 317  431  34.72      0.0      10  275   64  392    3
290
> 1mx1A      1  6303 4500  532  431  21.89      0.60E-02 11  141  112  254   49
217
> 1p0iA      1  6936 4650  523  431  27.52      0.26E-02 12   99  154  255  100
208

```

En este caso se creó un alineamiento objetivo y se le dio la instrucción a Modeller que leyera cada estructura PDB. Para ello fue necesario que todos los archivos PDB estuvieran en el mismo directorio que el programa. Se utilizó el método de **append\_model** para agregar cada estructura al alineamiento. Para mejorar el alineamiento se utiliza la herramienta **malign** para calcular un alineamiento de múltiples secuencias. El comando **malign3d** realiza una superposición de mínimos cuadrados de las cuatro estructuras usando el alineamiento de múltiples secuencias como punto de inicio.

```

from modeller import *
env = environ()
aln = alignment(env)
for (pdb, chain) in (('1u4n', 'A'), ('1jji', 'A'), ('1jkm', 'A'),
                    ('1lzl', 'A')):
    m = model(env, file=pdb, model_segment=('FIRST:'+chain, 'LAST:'+chain))
    aln.append_model(m, atom_files=pdb, align_codes=pdb+chain)
aln.malign()
aln.malign3d()
aln.compare_structures()
aln.id_table(matrix_file='family.mat')
    env.dendrogram(matrix_file='family.mat', cluster_cut=-1.0)

```

El comando **compare\_structures** compara las estructuras de acuerdo a los alineamientos construidos por **malign3d** y esto permite calcular diferencias entre los ángulos dihedrales de las cadenas laterales y principal, porcentajes de identidad de secuencias y varias otras mediciones. Finalmente con el comando **id\_table** se obtiene un archivo con parejas de distancias de secuencias que permiten utilizar el comando **dendrogram** que ayuda a la visualización de las diferencias entre los moldes candidatos. Esto se muestra en el siguiente archivo llamado **compare.log**.

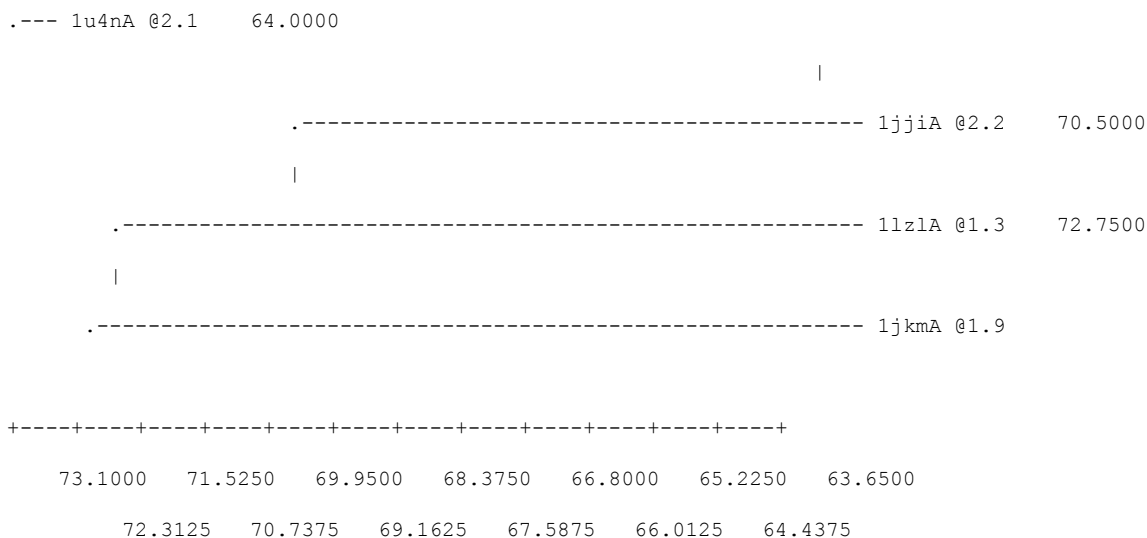
```

Sequence identity comparison (ID_TABLE):
    Diagonal      ... number of residues;
    Upper triangle ... number of identical residues;
    Lower triangle ... % sequence identity, id/min(length).

1u4nA @21jjiA @21jkmA @11lzlA @1
1u4nA @2      308      111      79      91
1jjiA @2      36       311     91      89
1jkmA @1      26       29     358     85
1lzlA @1      30       29     27     317

```

Weighted pair-group average clustering based on a distance matrix:



Finalmente para seleccionar el molde adecuado, se utilizó como forma complementaria los valores entregados por FUGUE. En el archivo fugue.out se entregan los valores de z-score de cada uno de las secuencias escogidas por FUGUE.

**Tabla I:** Tabla de valores de z-score de cada uno de los moldes.

# Profile	PLEN	RAWS	RVN	ZSCORE	PVZ	ZORI	EVP	EVF	AL#
hs1jjia	311	389	834	44.35	1.0E+03	46.85	1.0E+03	1.0E+03	00
hs2qrua	270	280	655	41.80	1.0E+03	44.06	1.0E+03	1.0E+03	02
hs1vkha	253	184	583	38.98	1.0E+03	41.44	1.0E+03	1.0E+03	02
hs3bjra	241	85	487	38.90	1.0E+03	41.34	1.0E+03	1.0E+03	02
hs2pbla	255	205	617	38.81	1.0E+03	41.19	1.0E+03	1.0E+03	02
hsd1jkma	358	247	718	35.62	1.0E+03	38.10	1.0E+03	1.0E+03	00
hs1nlma	726	-255	131	20.28	1.0E+03	22.91	1.0E+03	1.0E+03	33
hs2ecfa	700	-185	131	14.94	1.0E+03	17.36	1.0E+03	1.0E+03	33
hs2d5la	665	-222	120	14.55	1.0E+03	17.05	1.0E+03	1.0E+03	33
hs2fuka	218	-217	148	13.61	1.0E+03	16.08	1.0E+03	1.0E+03	02

A pesar de que todos los moldes presentan un valor de z-score sobre 6 (ver criterio en anexo 8.2.3), se seleccionó solo el molde 1jji para realizar el alineamiento ya que este molde presentó el mejor valor considerando todas las penalizaciones de gaps y sobre todo porque es el primero en el ranking que entregó FUGUE. Además, es la secuencia que presenta mayor número de aminoácidos alineados con la secuencia de la lipasa 2-17.

Cuando se comenzó el proceso de evaluación de los moldes, el programa Modeller arrojaba errores debido a que la similitud entre los modelos en el archivo compare.py era muy baja. Para que los modelos puedan ser comparados deben ser más o menos parecidos entre ellos. Cabe destacar que en esta parte, la comparación se hace independiente del molde. El archivo compare.py recibe sólo los PDB de los modelos que se utilizarán para hacer el modelamiento.

#### 2.4.2 Alineamiento de la lipasa 2-17 con el molde 1jjiA

Para alinear la secuencia de la lipasa 2-17 con la estructura de 1jjiA se utilizó el comando **align2d()** en Modeller. Aunque **align2d()** está basado en algoritmos de programación dinámica, este comando funciona de diferente forma para el método estándar de alineamiento de secuencias debido a que toma en cuenta información estructural a partir del molde para construir el alineamiento. Esto es realizado mediante una función de penalización de gaps que tiende a poner gaps en las regiones curvas y expuestas al solvente, fuera de los segmentos de estructura secundaria y entre dos posiciones que están muy cercanas. Como resultado, los errores en el alineamiento son reducidos aproximadamente a la mitad si lo comparamos con técnicas de alineación estándar.

En este caso, se alineó la secuencia de la lipasa 2-17 en el archivo Query.ali con el molde 1jjiA del archivo 1jjiA.pdb. Esto se puede apreciar en el archivo **align2d.py**.

```
from modeller import *
env = environ()
aln = alignment(env)
mdl = model(env, file='1jji', model_segment=('FIRST:A','LAST:A'))
aln.append_model(mdl, align_codes='1jjiA', atom_files='1jji.pdb')
aln.append(file='Query.ali', align_codes='Query')
aln.align2d()
aln.write(file='Query-1jjiA.ali', alignment_format='PIR')
aln.write(file='Query-1jjiA.pap', alignment_format='PAP')
```

Nuevamente se crea un archivo de entrada. Se creó un alineamiento vacío "aln" y luego un nuevo modelo de la proteína "mdl" en el cual se leyó la cadena A del archivo PDB 1jji. El comando **append\_model()** transfiere la secuencia del archivo PDB de este modelo de alineamiento y le asigna el nombre de " 1jjiA" (**align\_codes**). Posteriormente se agrega la secuencia de la lipasa desde el archivo

Query.sec al alineamiento usando el comando **append()**. El comando **align2d()** se utiliza entonces para alinear las dos secuencias. Finalmente el alineamiento es escrito en formato PIR (Quey-1ijiA.ali) y PAP (Quey-1ijiA.pap). El formato PIR es utilizado por Modeller en la etapa subsecuente de la construcción del modelo, mientras que el formato de alineamiento PAP permite una visualización más fácil. Como se observa en el archivo **Query-1ijiA.pap** (mostrado abajo) fue necesario que se incorporaran varios gaps ya que la similitud entre el molde y la lipasa 2-17 no era tan alta. Las posiciones con aminoácidos idénticos están marcados con (\*).

```

aln.pos      10      20      30      40      50      60
1jjiA      MLDMPIDPVYYQLAEYFDSLPK-----FD-----QF--SSAREYREA-
Query      MPILPVPSLNALLTKTLKTVKKTAAKNVHQHHILHHTLKGLDNLPTLLERVNHQLKASTAEQYPLAD
_consrvd   *  *      *      *      *      *      *      *      *

```

```

_aln.p      70      80      90      100     110     120     130
1jjiA      -----INRIYEERNRQLSQHE----R-----V-----E-----RVEDRTIKGRN-
Query      AHLRLILAVSNKLRPLALDKLPNLRQKFGTDAVSLQAPSVWQQNAEASGENSENAVSWQDKTIANADG
_consrvd   *      *  *      *      *      *      *      *  *

```

```

_aln.pos    140     150     160     170     180     190     200
1jjiA      GDIRVRVYQQ-----KP-DSPVLVYYHGGGFVICSIESHDALCRRRIARLSNSTVVSVDYRLAPEHK
Query      GDMTVRCYQQSKQNNKSTDEVAMLFFHGGGFCIGDIDTHHEFCHTVCAQGTGWAVVSVDYRMAPEYP
_consrvd   **  *  ***      *  *      *  *  *  *  *      *  *  *  *  *

```

```

_aln.pos    210     220     230     240     250     260     270
1jjiA      FPAAVYDCYDATKWVAENAEELRIDPSKIFVGGDSAGGNLAAAVS---IMARDSG-----E-D--
Query      APTALKDCLSAYAWLTEHSQSLGALPSRIVLSGDSAGGCLAALVAQQVIKPIDALWQDNNQATETDKK
_consrvd   *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *

```

```

_aln.pos    280     290     300     310     320     330     340
1jjiA      ----FIK-----HQILIYPVNVFVAPTPSLLEFGGLWILDQKIMSWFSEQYFSREE-DKFNPL
Query      ANDTFKKSADLPRPLAQLPLYPVTDYEAEPYSEWELYGGL-LLDHNDANVFNTAYTQHSGMLQSHPL
_consrvd   *  *      *  *  *  *  *  *  *  *  *  *  *  *  *

```

```

aln.pos      350      360      370      380      390      400
1jjiA      ASVIFADLENLPPALIIITAEYDPLRDEGEVFGQMLRRAGVEASIVRYRGVHLHGFINIYYPVLKAARDAI
Query      ISVMHGDNAHLCPYSYVVVAELDILRDEGLAYAELLQKEGVKVVETVTVLGAPHGFINLMSVHQGLGNQT
_consrvd   **   *   * *   ** * *****   *   **   *   *****   *

aln.p      410      420      430
1jjiA      NQIAALLV-F-----D----
Query      NYIIKSFACFVQNLLTSEGDEPNL
_consrvd   * *   *   *

```

Dado que la carboxilesterasa 1iji posee solo 315 aminoácidos, hubo muchas zonas con gaps, que quedaron sin molde, pero al analizar los otros alineamientos para complementar estas zonas no se encontró alguno que pudiese ser incluido para la generación de un alineamiento múltiple. De todas formas la evaluación del alineamiento fue positiva, ya que no se reportaron distancias atómicas de enlace entre carbonos  $\alpha$  equivalentes entre el molde y la secuencia blanco, con valores mayores que 6Å. Además el alineamiento presenta un porcentaje de identidad aceptado por Modeller.

### 2.4.3 Construcción del modelo

Una vez que se construye el alineamiento entre el molde y la proteína target, el programa Modeller calcula un modelo tridimensional de la proteína *target* automáticamente, usando **automodel**. Esto generó cinco modelos similares de la lipasa 2-17, basados en el molde 1jjiA, como se ve en el archivo **model-single.py**.

```

from modeller import *
from modeller.automodel import *
env = environ()
a = automodel(env, alnfile='Query-1jjiA.ali',
knowns='1jjiA', sequence='Query',
               assess_methods=(assess.DOPE, assess.GA341))
a.starting_model = 1
a.ending_model = 5
a.make()

```



La primera línea carga el **automodel** y lo prepara para su uso. Luego, se crea el objeto **automodel**, llamado “**a**”, y un set de parámetros para guiar la construcción del modelo. **alnfile** nombra los archivos que contienen el alineamiento entre el molde y el target en el formato PIR. **knowns** define las estructuras de moldes conocidas en **alnfile**. **starting\_modelending\_model** define el número de modelos que son calculados. El principal archivo generado por Modeller es el archivo **model-single.log** (mostrado abajo), el cual reporta advertencias o errores y otra información útil para el modelo final.

```
>> Summary of successfully produced models:
Filename                molpdf      DOPE score   GA341 score
-----
Query.B99990001.pdb    3622.22290  -40895.56250  1.00000
Query.B99990002.pdb    3373.26294  -40743.11719  1.00000
Query.B99990003.pdb    2988.39062  -41212.42188  1.00000
Query.B99990004.pdb    2959.33521  -40887.82422  1.00000
Query.B99990005.pdb    2982.06201  -40610.10156  1.00000
```

En este archivo se observa un resumen de los primeros cinco modelos construidos. Para cada modelo, se obtiene una lista con el nombre del archivo que contiene las coordenadas del modelo en formato PDB. Los modelos pueden ser analizados mediante cualquier programa que lea archivos PDB. En este caso se utilizó el programa “pdb viewer”. Estos archivos también entregan puntajes de cada modelo.

Se escogió trabajar con el modelo tres ya que presentó un DOPE score de -41212.42188. El DOPE (*Discrete Optimized Protein Energy*) es una herramienta estadística que mide la similitud con el molde. Es utilizado para determinar la energía del modelo proteico, generado mediante las diversas interacciones creadas por Modeller. Mientras más bajos son los puntajes entregados por el DOPE, mejor evaluado estará el modelo. Además, el DOPE puede generar perfiles de energía entre cada residuo del modelo, permitiendo localizar regiones problemáticas en la estructura del modelo establecido.

#### 2.4.4 Evaluación del modelo utilizando Verify 3D

Cuando se obtienen varios modelos determinados a partir del mismo molde, el mejor modelo puede ser evaluado de diferentes formas. Una de ellas es escogiendo el modelo con el valor más bajo entregado por el DOPE, o con el puntaje más alto entregado por GA341. La función **molpdf** permite ver la calidad de los átomos seleccionados en el modelo, y siempre es calculada y reportada en cada archivo PDB generado.

Los valores de DOPE no son absolutos, es decir solo pueden ser usados para posicionar modelos calculados a partir del mismo alineamiento. GA341 utiliza el porcentaje de identidad de secuencia entre el molde y el modelo como parámetro y los puntajes entregados fluctúan entre 0.0 para el peor modelo y 1.0 para el mejor, pero no es tan buena herramienta como DOPE para distinguir entre un buen y mal modelo. Una vez que un modelo es seleccionado, este debe ser evaluado. Antes de cualquier evaluación externa del modelo, se debe comprobar en el archivo log (model-single.log) si es que hay errores en los tiempos de corrida del modelo y violaciones en las restricciones. El archivo **evaluate\_model.py** evalúa el modelo de entrada con el potencial DOPE.

Para continuar la evaluación se pueden utilizar programas como Procheck o Verify 3D que permiten tener una mejor visión del modelo. En este caso se procedió a utilizar el programa Verify 3D que permite analizar la calidad de la estructura tridimensional. Para ello, se cargó el archivo **Query.B99990003.pdb**, correspondiente al modelo 3, en el programa Verify 3D el cual entregó un análisis para cada aminoácido (anexo 8.2.4).

Así, fue posible determinar visualmente qué aminoácidos se alejan del valor promedio 3D-1D. En este caso, la mayoría de los aminoácidos con mayores errores fueron justamente lo que no tuvieron una estructura secundaria atribuida.

#### **2.4.5 Evaluación preliminar de modelo mediante el servidor VADAR**

Una vez obtenido el modelo de la lipasa 2-17 se llevó a cabo una evaluación del modelo utilizando el servidor web VADAR (Willard y col., 2003), el cual señala posibles errores que pueden presentar los distintos residuos del modelo. VADAR (*Volume, Area, Dihedral Angle Reporter*) es un conjunto de muchos algoritmos diferentes y programas, para el análisis y la evaluación de estructuras de proteínas en base a sus archivos PDB (Willard y col., 2003). Los resultados han sido validados a través de la comparación extensiva de datos publicados, además de una cuidadosa inspección visual. El servidor VADAR entrega una serie de tablas y gráficos de alta calidad que permiten la evaluación cuantitativa y cualitativamente de estructuras de las proteínas.

La evaluación del modelo es una etapa fundamental, ya que ésta permite identificar los aminoácidos con posibles problemas estéricos, los cuales pueden ser corregidos mediante herramientas del software utilizado. En base a lo anterior, se sometió el modelo generado a una evaluación de su área superficial accesible, volumen fraccional, ángulos phi, psi y omega. En la Figura 6 se muestra el análisis de calidad de la primera estructura obtenida con Modeller, junto con la identificación de los residuos con posibles problemas. Los resultados entregados por el análisis mediante el servidor VADAR fueron claves para la determinación de los residuos que presentan problemas espaciales. De este modo, se

identificaron fundamentalmente los residuos que poseen problemas con su volumen fraccional y problemas en sus ángulos (Figura 6). El volumen fraccional es calculado como el volumen ocupado por un residuo definido por su radio atómico y sus vecinos más cercanos (Willard y col., 2003).

RES. NUM.	RES. NAME	SCND STRUC	HBOND HBOND	BTURN BTURN	RES. ASA	FRAC. ASA	RES. VOL.	FRAC. VOL.	PHI PHI	PSI PSI	OMEGA OMEGA	PRBLM PRBLM
3	ILE	CCC	C		74.7	0.38	143.3	0.89	-44.1	152.7	-169.2	O
6	VAL	CCC	C 8		41.9	0.24	114.6	0.84	-143.0	155.4	-169.3	O
8	SER	CCH	C 6,11	III	70.3	0.53	78.3	0.86	80.0	-46.6	-179.7	P
10	ASN	HHH	H 13	III	11.8	0.07	186.6	1.59	-61.8	-39.3	-177.9	V
19	THR	HHH	H 15,16	III	33.6	0.22	94.8	0.81	-69.6	-22.7	-166.1	O
21	LYS	CCC	C		55.8	0.26	150.3	0.98	117.9	114.5	-173.6	P
22	LYS	CBC	C		50.9	0.24	197.6	1.28	-148.9	-144.2	-178.9	V
42	ASP	CCC	C	III	21.8	0.14	123.3	1.08	-69.3	-46.0	-169.8	O
43	ASN	CBC	C 40	III	28.9	0.17	141.4	1.21	-84.4	105.7	174.6	V
48	LEU	CCC	C		162.2	0.78	136.8	0.84	69.0	124.2	-175.2	P
54	HIS	CCC	C		101.9	0.50	191.2	1.23	-124.2	-160.4	-177.9	V
56	LEU	CCC	C		5.8	0.03	129.6	0.79	-105.0	134.1	179.2	V
61	ALA	HHH	H 64,65		22.6	0.18	137.6	1.58	-64.5	-34.0	-173.7	V
66	LEU	HHH	H 62,76		127.6	0.61	148.4	0.91	-66.2	-61.9	-168.6	O
67	ALA	HHC	H 63		51.0	0.41	82.0	0.94	-70.7	-48.9	-163.9	O
72	ARG	CCC	C		30.6	0.13	246.6	1.41	-84.4	-152.2	-168.3	V,O
73	LEU	CCC	C		66.5	0.32	311.4	1.91	-91.7	118.3	169.8	V,O
74	ILE	CCC	C 77,78		0.0	0.00	122.2	0.76	-153.9	151.6	-166.8	V,O
77	VAL	HHH	H 81,80		4.8	0.03	98.3	0.72	-67.2	-42.8	-172.7	V
90	LEU	CCC	C		51.1	0.25	130.5	0.80	-139.5	-60.9	172.9	V
93	LEU	CCC	C		2.8	0.01	122.5	0.75	-95.5	-118.0	178.3	V
102	VAL	BCC	C		0.0	0.00	101.3	0.75	-23.5	130.0	-177.0	V
123	ALA	CCC	C 120	III	23.9	0.19	111.6	1.28	176.0	159.9	178.2	V
124	TRP	BBB	B		61.1	0.23	184.6	0.80	-109.9	135.3	178.8	V
127	GLN	BBB	B 143,143		55.5	0.29	175.2	1.26	-123.3	131.7	178.3	V
136	GLY	CCC	C	III	27.4	0.30	53.9	0.86	-68.0	-28.0	-168.5	O
143	CYS	BBB	B 127,127		2.4	0.02	131.7	1.26	-106.0	121.4	178.3	V
145	GLN	BBB	B 125		14.3	0.08	196.2	1.41	-117.8	134.6	-162.2	V,O
152	GLU	CCC	C 148		34.8	0.18	168.5	1.26	-53.8	147.8	-178.0	V
155	SER	CCC	C		17.1	0.13	117.8	1.30	-102.8	-178.1	175.6	V
156	THR	CCC	C 227		96.4	0.64	111.3	0.95	68.2	-40.2	-178.2	P
158	GLU	BBB	B 229		6.1	0.03	164.2	1.23	-130.4	159.9	178.3	V
160	ALA	BBB	B 234,232		1.7	0.01	111.7	1.28	-132.3	146.7	174.7	V
167	GLY	CCC	C 171,170	I'	18.6	0.20	87.2	1.38	118.7	13.6	171.8	V
168	GLY	CCC	C 170	I'	14.7	0.16	91.6	1.45	73.4	13.3	-179.0	V
170	CYS	CCC	C 167,168	I'	2.3	0.02	125.5	1.20	-95.0	-42.3	-166.5	V,O
172	GLY	CCC	C		5.8	0.06	82.4	1.31	87.0	169.2	-179.3	V
177	HIS	HHH	H 174,181		12.4	0.06	191.6	1.23	-93.7	0.5	-170.9	V
179	GLU	HHH	H 182,183		3.4	0.02	172.2	1.29	-61.6	-41.8	-176.7	V
191	ALA	BBB	B 159,161		2.2	0.02	109.3	1.25	-131.2	149.2	174.9	V
194	SER	BBB	B 142,142		0.0	0.00	122.0	1.35	-98.2	133.2	-173.1	V
200	ALA	CCC	C 169		0.3	0.00	84.8	0.97	-86.0	163.3	-1.0	C
205	ALA	CBC	C 327		5.4	0.04	103.0	1.18	-56.2	150.3	-1.6	C
206	PRO	HHH	H 209,210		47.4	0.31	126.7	1.10	-85.2	17.1	-169.5	V
223	SER	HHH	H 226,227	III	10.6	0.08	123.0	1.36	-62.0	-42.9	-175.5	V
225	SER	HHC	H 222	III	63.4	0.48	125.6	1.39	-66.7	-40.6	179.8	V
228	ALA	CCC	C		9.9	0.08	123.8	1.42	-124.5	157.2	-177.6	V
230	PRO	CCC	C	III	45.0	0.34	148.6	1.30	-58.9	-24.7	-178.9	V
238	ASP	BBB	B 164,241		6.8	0.04	142.8	1.26	-115.5	135.0	173.8	V
239	SER	CCH	C 293,243		15.1	0.11	130.2	1.44	72.6	-123.3	178.3	V
249	ALA	BHC	C 245		0.0	0.00	87.4	1.00	-82.2	-57.5	-166.8	O
250	GLN	CCC	C 246		0.7	0.00	127.5	0.92	-53.7	-158.8	-165.3	O
251	GLN	CCC	C		0.0	0.00	136.4	0.98	-46.6	-48.0	166.2	O
259	LEU	CHC	C 256	III	32.0	0.15	119.9	0.73	-65.7	-40.3	-173.4	V
266	ALA	CCC	C		83.2	0.51	167.7	0.78	-69.9	148.1	-175.9	O
267	THR	CCB	C 258		67.9	0.45	99.2	0.85	-110.5	-105.3	166.2	O
276	THR	CBC	C		0.8	0.01	104.4	0.89	-112.8	124.6	-166.4	O
278	LYS	CCB	C 270		0.0	0.00	146.0	0.95	-85.5	130.9	-159.3	O
280	SER	CCB	C 352		4.1	0.03	109.0	1.20	-90.1	139.1	-177.5	V
283	ASP	CCC	C		72.1	0.46	91.0	0.80	-77.4	72.5	176.1	V
288	LEU	BCB	B		1.8	0.01	134.4	0.82	-82.6	-141.8	-168.0	O
289	ALA	BCB	B 235		10.9	0.09	84.9	0.97	169.6	-149.8	177.3	P
295	PRO	CCC	C		1.6	0.01	145.4	1.36	-80.6	162.3	170.2	V
296	VAL	CCC	C		19.4	0.11	170.1	1.25	-107.3	109.7	178.6	V
297	THR	CCC	C 342,343		0.9	0.01	135.2	1.16	-95.8	-17.6	-168.5	O
299	TYR	CCC	C		29.8	0.12	246.0	1.29	-87.5	-32.9	-173.0	V
300	GLU	CCC	C		94.6	0.50	129.0	0.97	-77.2	-55.7	-169.3	O
313	LEU	CCC	C		29.4	0.14	198.9	1.22	-119.8	-175.9	175.0	V
314	LEU	CCC	C		59.6	0.29	267.1	1.64	-62.4	-32.9	177.1	V
316	ASP	CCC	C 319,320		20.1	0.13	148.7	1.31	-130.9	172.2	-173.8	V
317	ASP	HHH	H 323,322		11.5	0.07	156.4	1.37	-67.5	-34.9	-179.2	V
320	ALA	HHH	H 316,324		30.0	0.24	170.3	1.95	-62.9	-39.6	-179.2	V
324	ASN	HHH	H 320,328		24.7	0.15	164.0	1.40	-62.2	-37.7	-179.1	V
336	SER	CHH	H 332,333		39.7	0.30	110.9	1.22	-72.3	-16.6	-172.7	V
363	LEU	CHH	H 366,367	III	12.1	0.06	171.8	1.05	-71.4	-13.1	-169.6	O
366	GLU	HHH	H 363,370		0.3	0.00	160.0	1.20	-64.2	-40.2	-178.5	V
367	GLY	HHH	H 363,371		0.2	0.00	81.4	1.29	-62.8	-41.8	-178.7	V
390	PRO	CCC	C 393	II	18.7	0.12	159.3	1.38	-95.4	176.8	176.4	V
391	HIS	CCC	C	II	16.5	0.08	226.0	1.45	-62.2	142.0	175.7	V
392	GLY	CCC	C 395	II	16.9	0.19	111.0	1.76	80.1	17.8	-176.2	V
393	PHE	CCC	C 390,396	II	2.0	0.01	187.2	0.96	-65.4	-35.3	-169.7	O
395	ASN	CCC	C 392	I	33.5	0.20	246.2	2.10	-64.8	-24.6	-177.7	V
396	LEU	CCC	C 393,400	I	30.6	0.15	205.0	1.26	-73.0	2.4	-169.1	V,O
398	SER	HHC	H		98.9	0.75	116.1	1.28	-59.8	-43.1	-179.8	V
402	GLY	HHH	H 405,406		13.5	0.15	78.0	1.24	-57.8	-36.9	-179.1	V
403	LEU	HHH	H 400,407		0.1	0.00	130.1	0.80	-68.2	-38.4	-177.2	V
406	GLN	HHH	H 402,409		2.7	0.01	181.7	1.31	-60.7	-40.4	-179.2	V
411	ILE	HHH	H 407,415		10.2	0.05	124.3	0.77	-63.4	-38.7	-175.9	V
413	SER	HHH	H 416,410	III	27.1	0.21	117.5	1.30	-65.4	-38.7	-177.2	V
417	PHE	CCC	C		69.6	0.31	176.0	0.90	104.2	159.8	177.0	P
424	SER	CBC	C		15.1	0.11	133.0	1.47	-125.4	124.6	-179.4	V
426	GLY	CCC	C		34.2	0.38	87.5	1.39	130.4	179.6	-178.3	V

**Figura 6: Evaluación del modelo tridimensional de la lipasa 2-17 previo al proceso de refinamiento, realizada mediante la herramienta VADAR.** Los errores que se pueden identificar con el servidor utilizado son: **A:** Indica posible problema con ASA4 fraccional (fASA > 1.0). **V:** Indica posible problema con volumen fraccional (fV < 0.8 o fV > 1.2). **P:** Indica posible problema con ángulos Phi y Psi (Morris y col., 1992). **O:** Indica posible problema con ángulo omega (omega > 170 u omega < -170) **C:** Indica enlace peptídico cis (-20 < omega < 20).

Normalmente, si la proteína se plegó correctamente, sus residuos debiesen presentar volúmenes fraccionales cercanos a 1. Valores menores a 0.8, indican que dichos residuos se ubican en una región comprimida. Por otro lado, el ángulo omega mide el giro en torno al enlace peptídico. En el caso de los residuos indicados en la tabla con la letra O, éstos poseen su ángulo omega fuera de rango (valores menores a 170 o mayores a -170), lo cual indica posibles problemas con la conformación *trans* de los enlaces. En el modelo sin refinar fue posible establecer que 92 aminoácidos presentaban problemas, dentro de los cuales, 65 presentaron problemas en el volumen fraccional y solo 25 en el ángulo omega. Particularmente, 29 aminoácidos corresponden a aminoácidos presentes en una estructura  $\alpha$  hélice, 7 presentes en hojas  $\beta$  y 56 a aminoácidos en *loops*.

#### **2.4.6 Refinamiento con Modeller**

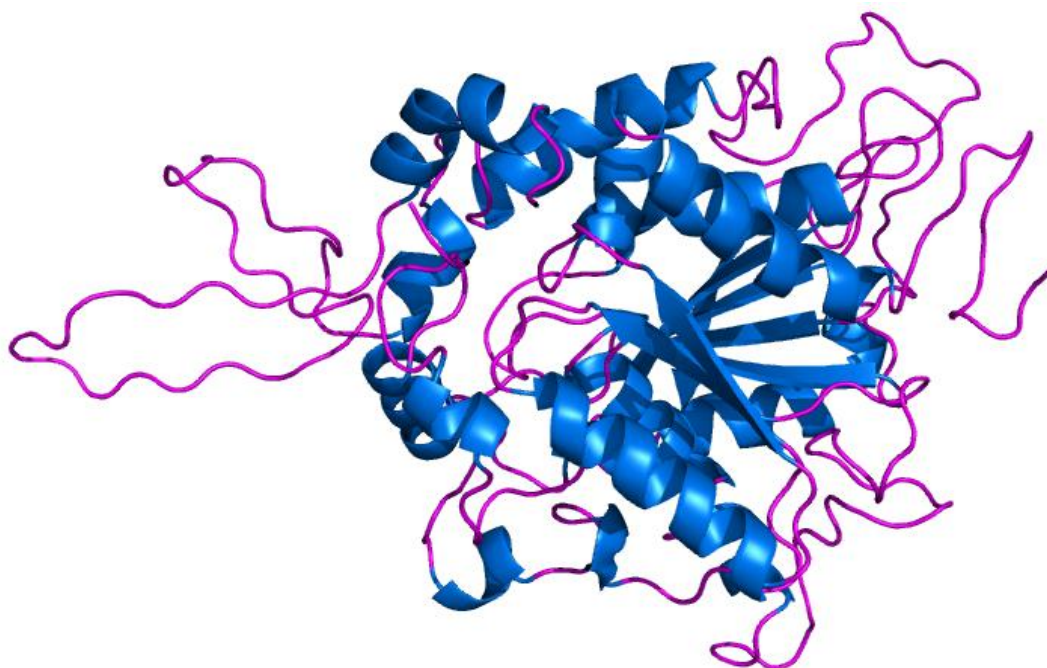
Una vez realizada la evaluación mediante VADAR, se llevó a cabo el refinamiento del modelo tridimensional obtenido para la lipasa 2-17 utilizando la herramienta de refinamiento de Modeller, con el objetivo de mejorar principalmente los *loops* que contaban con el mayor número de aminoácidos problemáticos. Modeller lleva a cabo el refinamiento utilizando dinámica molecular con el algoritmo probabilístico SA (*simulated annealing*) (Sali & Blundell, 1994), un método que permite encontrar el mínimo global de una función de costos que puede poseer varios mínimos locales.

Una vez llevado a cabo el refinamiento con Modeller es posible generar el perfil DOPE del modelo antes y después de ser refinado. Como se mencionó, el perfil DOPE corresponde a una función de energía y permite identificar zonas del modelo con posibles errores (Shen & Sali, 2006).

Para comenzar el refinamiento de la lipasa 2-17, lo primero fue establecer los aminoácidos presentes en los *loops* ya que éstos no presentan una estructura tridimensional bien definida y podrían corresponder a estructuras hélices  $\alpha$  o bien hebras  $\beta$  que no fueron modeladas debido a la falta de similitud con la estructura molde, además de presentar la mayor cantidad de problemas según la evaluación entregada por VADAR.

##### **2.4.6.1 Selección de los aminoácidos para refinamiento**

Para comenzar el refinamiento, utilizando el programa Pymol se seleccionaron los aminoácidos que se encontraban en los principales *loops* que presentaron problemas, particularmente, desde metionina 1 hasta leucina 9, desde valina 20 hasta treonina 60, desde lisina 89 hasta serina 124, desde glutamina 144 hasta ácido glutámico 157, desde fenilalanina 163 hasta isoleucina 173, desde tirosina 196 hasta treonina 206, desde triptófano 259 hasta alanina 288, desde tirosina 293 hasta prolina 303, desde tirosina 326 hasta glicina 331 y finalmente desde cisteína 415 hasta leucina 430.



**Figura 7: Distribución de aminoácidos en la estructura secundaria de la lipasa 2-17.** En azul se muestran regiones que tienen una estructura secundaria atribuida tipo  $\alpha$  hélice o bien hoja  $\beta$ . En púrpura se encuentran indicadas las regiones que forman *loops*.

Para cada uno de los casos, los aminoácidos fueron seleccionados, coloreados y evaluados en la estructura tridimensional creada previamente. Para determinar su ubicación espacial se seleccionaron los aminoácidos de las estructuras  $\alpha$  hélices y hojas  $\beta$ .

Finalmente se escogieron 10 *loops* los que fueron refinados utilizando Modeller (Figura 7). Cada *loop* fue refinado de manera independiente para realizar las evaluaciones. De la misma forma, se hizo un refinamiento de cada *loop* utilizando como molde el refinamiento anterior, esto es, el primer refinamiento se utilizó como molde para refinar un nuevo *loop*, y así sucesivamente.

#### **2.4.6.2 Refinamiento aminoácidos 259-288**

Para el caso del primer refinamiento, correspondiente al *loop* formado por los aminoácidos desde el triptofano 259 hasta la alanina 288, Modeller calculó 10 posibles conformaciones para el *loop* que permitieran mejorar las interacciones de los diferentes aminoácidos. El programa Modeller crea archivos .pdb para visualizar cada una de estas nuevas conformaciones.

Así, para comenzar el refinamiento, se le indicó a Modeller que creara 10 modelos con las posibles nuevas conformaciones de estos aminoácidos. Modeller entregó 10 nuevos archivos en formato pdb que fueron analizados con programas como Pymol. Además Modeller entregó un archivo con el análisis de cada modelo y una tabla final indicando los valores molpdf de cada modelo.

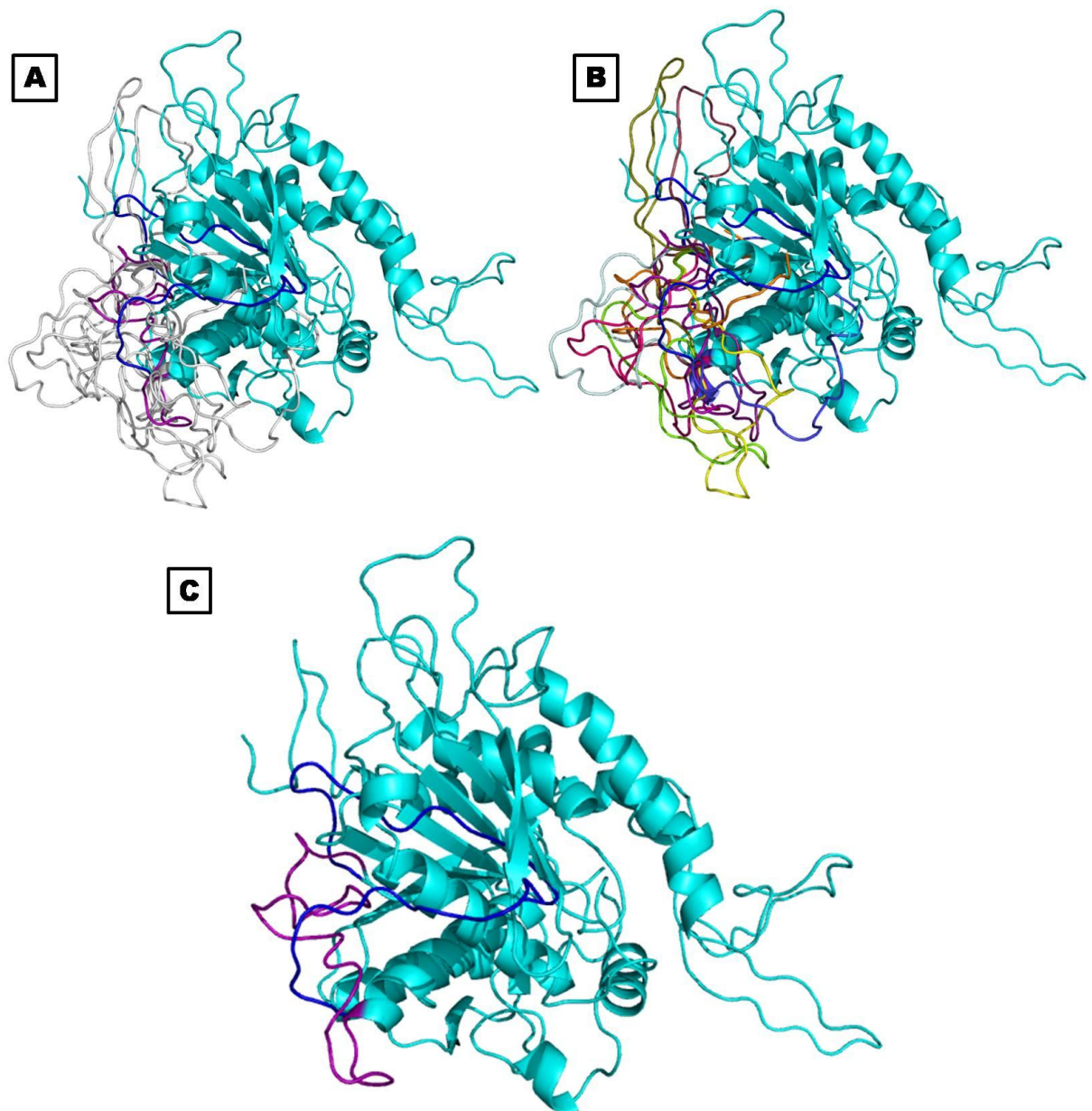
Como se mencionó previamente, el comando molpdf evalúa la calidad de los átomos seleccionados en el modelo usando el método DOPE (Shen & Sali, 2006), que corresponde a un potencial estadístico optimizado para el modelo evaluado. Se seleccionó el valor más alto de molpdf para cada uno de los refinamientos, el que según las evaluaciones realizadas por Modeller, resultó ser el modelo con la mejor conformación espacial para cada uno de los aminoácidos que conforman el *loop* (Figura 8).

<b>Aminoácidos: 259-288</b>	
<b>Nombre del archivo</b>	<b>molpdf</b>
Query.BL00010001.pdb	58,26345
Query.BL00020001.pdb	78,38163
Query.BL00030001.pdb	62,06118
Query.BL00040001.pdb	4860,33398
Query.BL00050001.pdb	6966,42676
Query.BL00060001.pdb	5383,28516
Query.BL00070001.pdb	81,55203
Query.BL00080001.pdb	1043,7439
Query.BL00090001.pdb	3490,44653
Query.BL00100001.pdb	145,22685

**Figura 8: Valores molpdf entregados por Modeller al refinar el *loop* 259-288.** En el cuadrado rojo se destaca el modelo mejor evaluado.

Modeller entregó para cada uno de los refinamientos un archivo pdb que permitió comparar la distribución espacial de cada uno de los aminoácidos y visualizar los modelos antes y después del refinamiento. En la Figura 9 es posible comparar el ordenamiento espacial del *loop* formado por los aminoácidos desde el triptófano 259 hasta la alanina 288, antes y después del refinamiento.

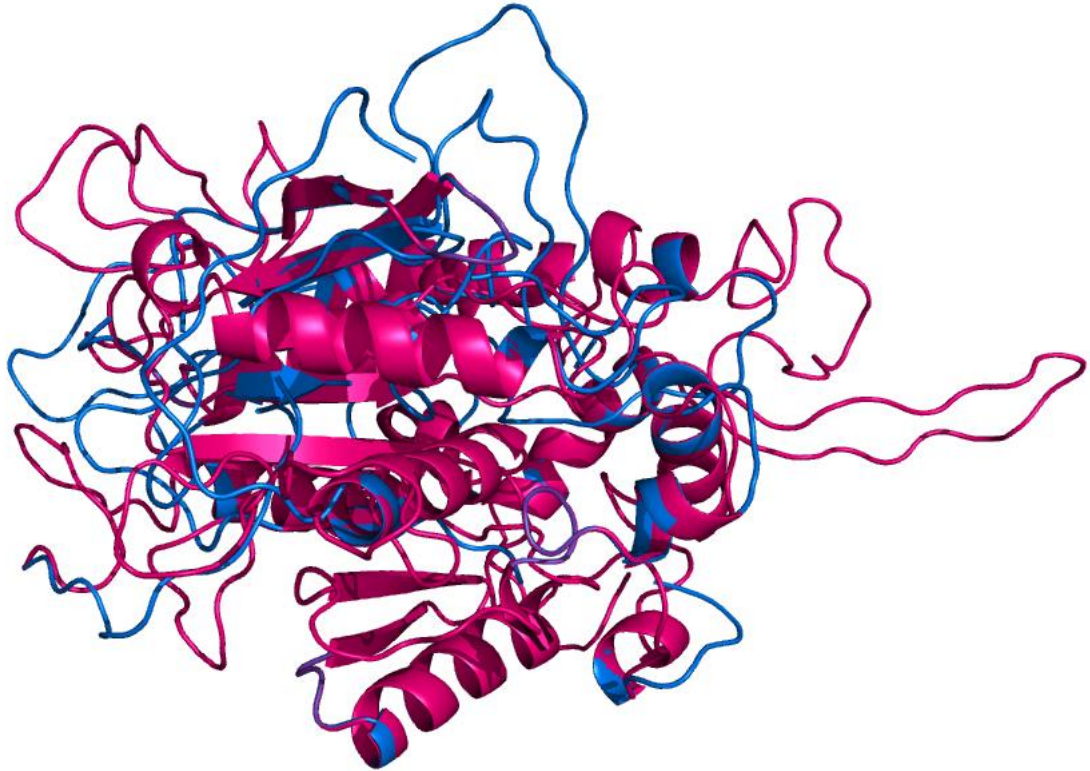
El *loop* una vez refinado (en azul) exhibe un ordenamiento más compacto en comparación al *loop* sin refinar (en púrpura). Además es posible determinar en el modelo que el nuevo *loop* tiene más interacciones con otros aminoácidos presentes en estructuras secundarias que el *loop* sin refinar. Debido a que las zonas de los *loops* son zonas con muy baja identidad con la proteína molde, en comparación a las estructuras  $\alpha$  hélices y hojas  $\beta$ , no es posible estimar cuál es el correcto posicionamiento de los átomos que lo conforman en base al modelo creado por homología. Esto puede provocar que Modeller fuerce el ordenamiento del nuevo *loop* desestabilizando nuevas regiones, lo que conlleva a una nueva conformación de la proteína completa, que a diferencia de lo esperado, no se ve favorecida energéticamente en comparación a la original. Para ello, todo nuevo refinamiento debe ser evaluado por herramientas como VADAR, que hacen un análisis de la proteína completa.



**Figura 9: Orientaciones para el *loop* formado desde el aminoácido 259 hasta el 288 una vez realizado el refinamiento. A:** Lipasa 2-17 con todas las posibles orientaciones del *loop* 259-288. En celeste es posible observar la parte de la proteína que no fue refinada. En colores, todas las posibles conformaciones del *loop* 259-288. **B:** Lipasa 2-17 con todas las orientaciones del *loop* destacándose en púrpura el *loop* inicial, en azul el *loop* final y mejor evaluado. En gris el resto de los *loops* que no fueron seleccionados. **C:** Modelo de la lipasa 2-17 que exhibe las dos conformaciones del *loop* 259-288. Púrpura: conformación inicial. Azul: conformación final una vez refinado.

Se repitió el mismo procedimiento para cada uno de los *loops* descritos en la sección 2.4.6.2, comparando el *loop* en el modelo sin refinar y el *loop* obtenido por el refinamiento mejor evaluado. Se compararon los valores entregados por Modeller y se escogió al mejor evaluado en cada caso (ver evaluación en anexo 8.2.5).

Una vez refinados cada uno de los *loops* se comparó el modelo inicial sin refinamiento y el modelo con cada uno de los *loops* refinados. Se alinearon utilizando el programa Pymol y se comparó la ubicación espacial de cada uno de ellos.

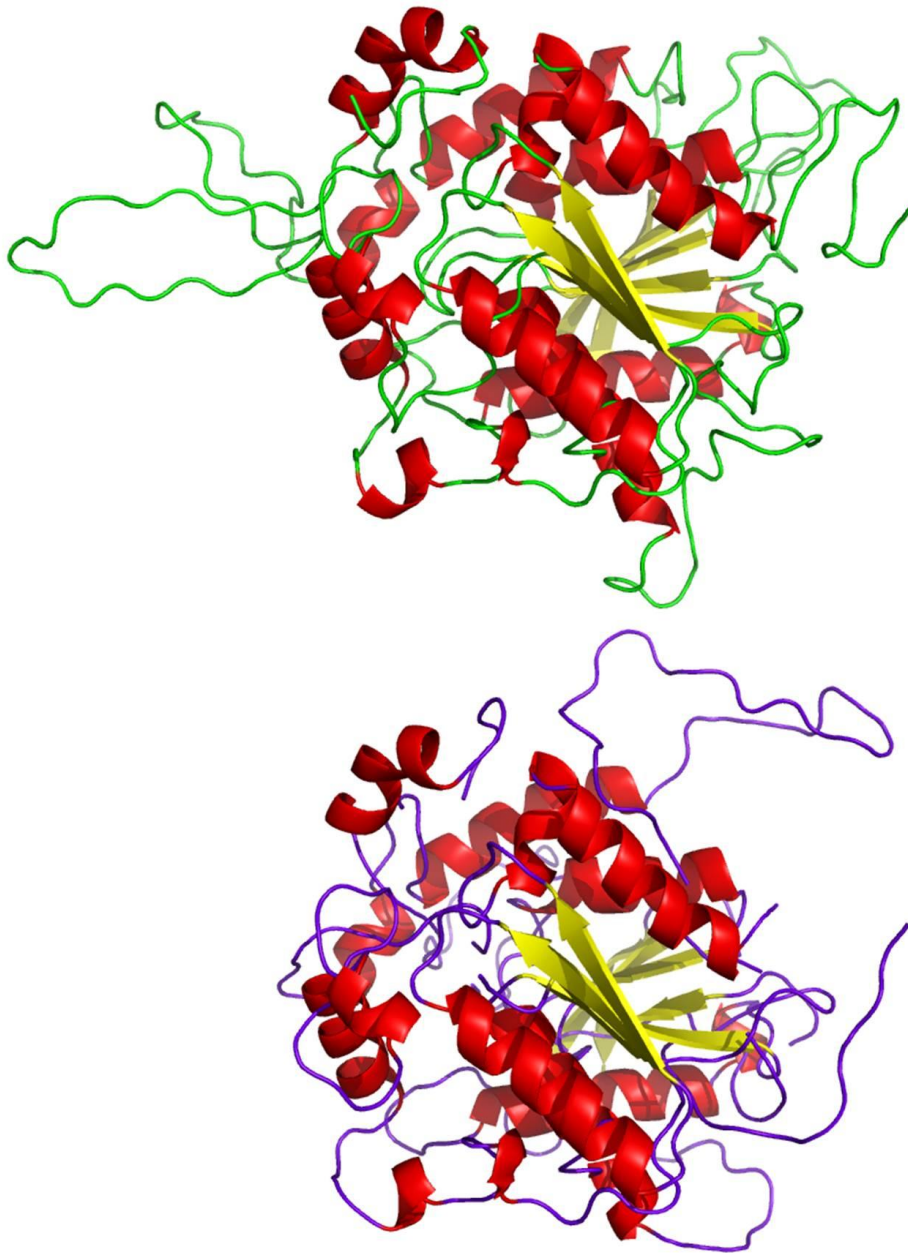


**Figura 10: Modelo refinado de la lipasa 2-17 superpuesto con el modelo sin refinar de la lipasa 2-17.** En púrpura se muestra el modelo inicial y en celeste el modelo final. Las estructuras secundarias  $\alpha$  hélices y hojas  $\beta$  se encuentran superpuestas y exhiben una distribución espacial muy similar, mientras que los *loops* presentan un nuevo posicionamiento.

Las estructuras de las lipasas, previo y posterior al refinamiento, fueron alineadas en base a la estructura principal tipo barril mientras que los *loops* quedaron libres para comparar las diferencias espaciales entre ambas estructuras (Figura 10).

A pesar de alinear la estructura tipo barril, es posible apreciar diferencias entre la posición de ambas hojas  $\beta$  y de algunas  $\alpha$  hélices entre los dos modelos. Esto es debido a que Modeller realizó modificaciones en otros aminoácidos que no pertenecían a los *loops* escogidos para el refinamiento y por lo tanto realizó un cambio estructural global en la proteína, que es apreciable principalmente en las estructuras  $\alpha$  hélices. En la Figura 11 es posible apreciar cómo cambia el ordenamiento de los *loops* en base al alineamiento de la estructura tipo barril, que a su vez presenta ciertas diferencias en las estructuras secundarias definidas.





**Figura 11: Comparación de modelos de la lipasa 2-17. Arriba:** Modelo sin refinar. **Abajo:** Modelo refinado. Para ambos casos el color rojo corresponde a las estructuras  $\alpha$  hélices, el color amarillo a las hojas  $\beta$ . El color verde y el púrpura corresponden a los *loops* antes y después del refinamiento, respectivamente.

#### 2.4.7 Evaluación mediante VADAR del modelo generado y refinado por Modeller

Una vez realizado el refinamiento del modelo estructural de la lipasa 2-17 fue necesario llevar a cabo una nueva evaluación, para determinar si los problemas detectados previamente fueron solucionados. En la Figura 12 se muestra el análisis de calidad del modelo de la lipasa 2-17 después de ser refinado, utilizando VADAR.

RES. NUM.	RES. NAME	SCND STRUC	HBOND HBOND	BTURN BTURN	RES. ASA	FRAC. ASA	RES. VOL.	FRAC. VOL.	PHI PHI	PSI PSI	OMEGA OMEGA	PRBLM PRBLM
2	PRO	CCC	C		0.0	0.00	68.6	0.60	-65.0	-78.9	155.9	V,O
8	SER	CCC	C		55.1	0.42	97.6	1.08	-163.9	-101.9	-171.7	P
9	LEU	CCC	C	III	67.5	0.33	149.4	0.92	-111.3	-166.5	-166.4	O
16	THR	HHH	H 13,19		2.7	0.02	92.3	0.79	-64.9	-40.3	-178.5	V
19	THR	HHH	H 15,16	III	14.7	0.10	77.8	0.67	-69.6	-20.0	162.1	V,O
21	LYS	CCC	C		15.1	0.07	154.9	1.01	-112.1	55.6	167.8	O
22	LYS	CCC	C 73		10.5	0.05	139.1	0.90	167.9	-123.5	179.1	P
23	THR	CCC	C		7.5	0.05	122.9	1.05	-118.3	75.0	157.7	O
25	ALA	CCC	C		36.5	0.30	108.0	1.24	-68.7	137.7	-172.3	V
32	HIS	CBG	C		94.0	0.46	136.9	0.88	-141.7	113.8	169.3	O
38	LEU	BBC	B		0.9	0.00	95.3	0.58	-112.0	-80.2	-172.5	V
40	GLY	CBG	C		1.7	0.02	78.1	1.24	-51.4	127.6	-173.1	V
42	ASP	CCC	C	II	25.1	0.16	143.0	1.26	-53.0	139.8	-171.2	V
47	PRO	CCC	C		47.5	0.31	106.8	0.93	-81.5	159.9	-164.4	O
48	LEU	CCC	C		1.7	0.01	131.3	0.80	-106.2	-114.9	-168.6	O
49	LEU	CCC	C		66.0	0.32	210.2	1.29	155.2	122.6	-175.5	V
50	GLU	CCC	C 53	III	9.0	0.05	168.8	1.27	-139.2	104.1	-172.6	V
55	GLN	BBB	B		3.0	0.02	149.6	1.08	-17.0	-158.3	171.8	P
66	LEU	HHH	H 62,76		48.2	0.23	149.4	0.91	-66.2	-61.9	-168.6	O
67	ALA	HHC	H 63		51.0	0.41	77.5	0.89	-70.7	-48.9	-163.9	O
69	ALA	CCC	C		46.9	0.38	69.4	0.80	-145.7	134.4	177.5	V
70	HIS	CCC	C		77.0	0.38	124.1	0.80	-157.1	-168.2	-179.0	V
71	LEU	CCC	C 3		29.6	0.14	126.8	0.78	-67.5	-42.5	-173.5	V
72	ARG	BBB	B		10.3	0.04	197.5	1.13	-84.4	-152.2	-168.3	O
73	LEU	BBB	B 22		2.4	0.01	112.0	0.69	-91.7	118.3	169.6	V,O
74	ILE	BBB	B 77,78		0.0	0.00	106.3	0.66	-153.9	151.6	-166.8	V,O
75	LEU	BHH	H 65,79		1.1	0.01	122.0	0.75	-62.9	-43.2	-173.3	V
77	VAL	HHH	H 81,80		4.8	0.03	100.0	0.74	-67.2	-42.8	-172.7	V
88	ASP	HHC	H 85	III	47.5	0.30	143.9	1.26	-71.5	-23.8	167.3	V,O
92	ASN	CBG	C		49.7	0.30	158.3	1.35	-79.6	153.7	-176.6	V
111	GLN	CBG	C		46.9	0.25	170.0	1.22	-108.5	-176.6	-173.7	V
117	SER	CCC	C		38.0	0.29	86.1	0.95	-13.6	-133.6	164.3	P,O
118	GLY	CCC	C		0.0	0.00	37.8	0.60	9.3	107.5	-170.3	V
119	ASN	CCC	C 161,184		0.0	0.00	67.2	0.57	-61.8	-66.5	-172.9	V
120	SER	CCC	C		0.0	0.00	45.5	0.50	127.7	69.3	-157.8	V,P,O
121	GLU	CCC	C		34.3	0.18	106.2	0.80	56.1	166.1	-134.9	V,P,O
122	ASN	CCC	C		38.8	0.23	109.6	0.93	-161.1	65.9	-163.7	O
131	ILE	BBB	B 139,139		0.0	0.00	100.5	0.62	-129.1	167.4	-176.6	V
133	ASN	BBB	B 273,137		4.6	0.03	78.8	0.67	-119.0	151.3	-175.9	V
136	GLY	CCC	C 198	III	8.4	0.09	45.6	0.72	-68.0	-28.0	-168.5	V,O
137	GLY	CCC	C 133,133	III	13.4	0.15	44.3	0.70	80.6	-117.4	-177.2	V
139	MET	BBB	B 131,131		0.3	0.00	81.7	0.50	-140.5	127.9	-179.2	V
141	VAL	BBB	B 129,129		0.0	0.00	93.2	0.69	-136.0	150.7	178.8	V
144	TYR	BCB	B 192		21.7	0.09	188.4	0.98	-106.0	-99.2	164.0	O
146	GLN	CCC	C		0.0	0.00	75.2	0.54	140.5	178.7	98.2	V,O
147	SER	CCC	C		0.0	0.00	37.5	0.41	132.7	101.6	-179.7	V,P
148	LYS	CCC	C		67.3	0.31	111.0	0.72	-165.2	144.1	170.9	V
149	GLN	CCC	C		0.0	0.00	84.1	0.61	-131.9	-84.2	170.3	V
150	ASN	CCC	C 231		22.5	0.14	147.4	1.26	135.2	91.8	-176.5	V,P
151	ASN	CCC	C		98.6	0.60	112.1	0.96	32.3	119.0	168.0	P,O
152	GLU	CCC	C 230		85.2	0.45	149.5	1.12	128.4	83.4	179.0	P
153	GLY	CCC	C		2.7	0.03	48.3	0.77	128.8	17.0	126.5	V,O
154	LYS	BCB	B		3.9	0.02	103.3	0.67	-112.6	-165.5	165.0	V,O
155	SER	BCB	B 230,228		21.4	0.16	87.9	0.97	-80.1	-131.2	167.5	O
158	GLU	BBB	B 229,227		3.0	0.02	87.4	0.66	-123.0	159.9	-178.3	V
159	VAL	BBB	B 191		0.0	0.00	60.4	0.44	-66.7	154.4	-178.8	V
160	ALA	BBB	B 234,232		0.0	0.00	42.6	0.49	-132.3	146.7	174.7	V
161	MET	BBB	B 191,193		0.0	0.00	124.5	0.76	-118.0	119.5	175.8	V
162	LEU	BCB	B 419,234		0.0	0.00	110.0	0.67	-83.7	122.2	-169.7	V,O
164	PHE	BCC	C		1.5	0.01	164.8	0.84	-102.4	-96.6	163.7	O
165	HIS	CCC	C		0.0	0.00	118.8	0.76	-73.7	-41.4	175.4	V
166	GLY	CCC	C		0.0	0.00	47.6	0.76	137.3	57.8	-176.2	V
167	GLY	CBB	B		0.0	0.00	43.8	0.70	-179.5	178.6	-128.0	V,O
168	GLY	CBB	B 198		0.3	0.00	40.4	0.64	-175.3	131.6	169.4	V,O
169	PHE	CBB	B 170		64.0	0.28	150.9	0.77	165.6	-177.7	-156.2	V,O
170	CYS	CCC	C 169		4.2	0.03	80.7	0.77	36.6	100.9	-170.3	V
171	ILE	BBB	B 46		5.2	0.03	114.3	0.71	-83.6	105.8	149.1	V,O
172	GLY	BBB	B 196		0.0	0.00	54.5	0.87	-48.8	133.5	-154.5	O
174	ILE	BHC	C 177	III	0.6	0.00	128.1	0.79	-98.7	-40.4	-177.6	V
179	GLU	HHH	H 182,183		13.6	0.07	184.9	1.39	-61.6	-41.8	-176.7	V
181	CYS	HHH	H 177,185		7.9	0.05	126.1	1.21	-66.8	-38.0	-178.6	V
184	VAL	HHH	H 180,119		0.0	0.00	99.0	0.73	-63.6	-42.7	-179.7	V
188	THR	CHH	H 185,417	III	0.0	0.00	67.5	0.58	-74.4	-31.4	-179.0	V
189	GLY	CCH	C 185,417	III	9.2	0.10	46.0	0.73	72.8	36.4	-179.0	V
190	TRP	BBB	B		2.6	0.01	122.2	0.53	-137.4	-141.3	-177.7	V
191	ALA	BBB	B 159,161		0.0	0.00	29.9	0.34	-131.2	149.2	174.9	V
192	VAL	BBB	B 144		1.3	0.01	107.9	0.79	-133.1	135.4	178.0	V
193	VAL	BBB	B 161,163		0.0	0.00	94.3	0.69	-116.5	111.6	-177.6	V
194	SER	BCB	B 142,142		1.7	0.01	98.5	1.09	-98.2	-167.5	143.8	O
195	VAL	BCB	B		0.0	0.00	87.9	0.65	164.9	-150.6	122.9	V,O
196	ASP	BCB	B 172		1.2	0.01	127.0	1.12	-111.0	114.9	152.6	O
197	TYR	BCB	B 215		0.0	0.00	118.6	0.62	-11.2	18.0	-128.9	V,O
198	ARG	CCC	C 168,136		44.7	0.18	140.9	0.80	139.4	-116.6	172.4	P
200	ALA	BCC	C		2.4	0.02	86.1	0.99	-123.7	174.9	-161.9	O
205	ALA	CCC	C		12.5	0.10	82.0	0.94	-89.8	-157.8	-150.4	O
206	PRO	BCC	C		18.3	0.12	129.8	1.13	-19.1	71.3	165.4	O
209	LEU	HHH	H 212,213		0.0	0.00	126.6	0.77	-58.0	-46.2	-176.0	V
211	ASP	HHH	H 207,215		0.0	0.00	79.5	0.70	-69.6	-32.6	177.6	V
212	CYS	HHH	H 208,216		0.0	0.00	80.9	0.77	-66.5	-41.9	179.9	V
215	ALA	HHH	H 219,211		0.0	0.00	64.3	0.74	-64.0	-41.4	-177.4	V
216	TYR	HHH	H 212,220		4.0	0.02	152.9	0.80	-63.9	-46.6	-179.2	V
219	LEU	HHH	H 215,216		0.0	0.00	98.9	0.61	-63.2	-45.5	-176.9	V
220	THR	HHH	H 216,217		26.1	0.17	144.4	1.24	-65.1	-40.2	-176.8	V
223	SER	HHH	H 226,227	III	6.7	0.05	72.4	0.80	-62.0	-42.9	-175.5	V
226	LEU	HHC	H 222,223	III	7.5	0.04	108.6	0.67	-68.0	-29.4	-179.7	V
228	ALA	BCB	B 155		0.0	0.00	57.9	0.66	-124.5	157.2	-177.6	V
229	LEU	BCB	B 158,232		11.1	0.05	122.7	0.75	-83.9	111.7	-178.9	V
232	ARG	BBB	B 229,160		3.1	0.01	130.0	0.74	-120.4	84.6	178.2	V

RES. NUM.	RES. NAME	SCND STRUC	HBOND HBOND	BTURN BTURN	RES. ASA	FRAC. ASA	RES. VOL.	FRAC. VOL.	PHI PHI	PSI PSI	OMEGA OMEGA	PRLM PRLM
233	ILE	BBB B	421		0.0	0.00	109.1	0.68	-115.7	126.5	-179.7	V
234	VAL	BBB B	160,162		0.0	0.00	92.5	0.68	-122.3	154.6	177.8	V
238	ASP	CBC C	241		8.3	0.05	148.5	1.31	-115.5	135.0	173.8	V
239	SER	CCH C	293,243		5.5	0.04	111.3	1.23	72.6	-123.3	178.3	V
240	ALA	HHH H	243,244		48.4	0.39	126.4	1.45	-60.0	-32.1	-179.7	V
243	CYS	HHH H	239,247		27.3	0.19	133.2	1.27	-61.4	-43.7	-177.2	V
246	ALA	HHH H	242,243		0.0	0.00	68.9	0.79	-64.1	-45.3	-177.1	V
247	LEU	BHH H	243,254		0.3	0.00	126.2	0.77	-67.2	-40.6	178.7	V
249	ALA	BHC C	245		2.9	0.02	103.6	1.19	-82.2	-57.5	-166.8	O
250	GLN	CCC C	246		0.0	0.00	119.8	0.86	-53.7	-158.8	-165.3	O
251	GLN	CCC C			0.0	0.00	139.1	1.00	-46.6	-48.0	166.2	O
265	GLN	CCC C			52.8	0.28	156.3	1.12	-116.3	134.1	167.2	O
273	ALA	CBC C	133		31.2	0.25	91.2	1.05	-171.3	87.8	-167.4	O
274	ASN	CCC C			12.1	0.07	95.6	0.82	-117.6	10.9	-146.1	O
275	ASP	CCC C			0.9	0.01	82.3	0.72	43.8	-65.3	87.6	V,P,O
276	THR	CCC C			0.0	0.00	69.1	0.59	-23.8	117.8	161.5	V,O
277	PHE	CCC C			52.0	0.23	146.3	0.75	-57.3	-111.6	141.0	V,O
278	LYS	CCC C			0.0	0.00	103.5	0.67	153.3	-171.8	178.5	V
280	SER	CCC C			0.0	0.00	62.2	0.69	52.6	93.2	-143.4	V,O
281	LEU	CCC C	232,418		0.0	0.00	87.6	0.54	-143.6	-174.3	153.0	V,O
282	ALA	CCC C			0.0	0.00	56.1	0.64	156.6	-171.4	163.3	V,O
283	ASP	CCC C			84.9	0.54	103.8	0.91	-113.2	-28.4	163.4	O
284	LEU	CCC C			98.7	0.48	145.8	0.89	176.2	60.7	161.0	O
286	ARG	BBC B			6.5	0.03	169.9	0.97	171.2	154.5	153.8	O
288	LEU	BBB B			6.6	0.03	126.0	0.77	-76.7	105.8	176.0	V
293	LEU	BBB B	237,239		10.8	0.05	190.0	1.16	-91.2	70.4	-163.9	O
294	TYR	BBB B	356		14.1	0.06	242.0	1.27	-84.2	132.1	-169.0	V,O
295	PRO	BBB B	367		2.6	0.02	105.9	0.92	-80.4	163.0	-165.5	O
296	VAL	BBB B	364		4.7	0.03	87.8	0.65	-92.0	118.2	155.8	V,O
297	THR	CCC C	309		0.0	0.00	66.5	0.57	25.1	107.4	-142.8	V,O
309	TYR	CCH H	297,305		19.6	0.08	152.0	0.79	-109.2	39.0	-178.5	V
327	TYR	HHH H	323,324		33.6	0.14	258.8	1.35	-62.5	-50.3	177.9	V
350	LEU	CBC C	347,348		1.9	0.01	128.2	0.79	-79.8	158.0	179.9	V
351	CYS	CBB B			2.6	0.02	128.3	1.23	-66.2	145.6	176.7	V
362	ILE	BHC C		III	3.2	0.02	114.8	0.71	-60.8	-40.8	-177.5	V
363	LEU	BHH H	366,367	III	36.4	0.18	140.7	0.86	-71.4	-13.1	-169.6	O
366	GLU	HHH H	363,370		35.3	0.19	175.6	1.32	-64.2	-40.2	-178.5	V
390	PRO	CCC C	393	II	20.8	0.13	149.8	1.30	-95.4	176.8	176.4	V
391	HIS	CCC C		II	4.6	0.02	192.1	1.23	-62.2	142.0	175.7	V
392	GLY	CCC C	395	II	10.2	0.11	95.1	1.51	80.1	17.8	-176.2	V
393	PHE	CCC C	390,396	II	7.4	0.03	203.6	1.04	-65.4	-35.3	-169.7	O
395	ASN	CCC C	392	I	4.3	0.03	182.4	1.56	-64.8	-24.6	-177.7	V
396	LEU	CCC C	393,400	I	14.1	0.07	189.8	1.16	-73.0	2.4	-169.1	O
397	MET	CHC C	394,115	I	12.9	0.06	122.7	0.75	-48.8	-37.1	-176.6	V
398	SER	BHC C	114		15.4	0.12	139.3	1.54	-59.8	-43.1	-179.8	V
402	GLY	HHH H	405,406		13.5	0.15	79.9	1.27	-57.8	-36.9	-179.1	V
403	LEU	HHH H	400,407		0.1	0.00	118.7	0.73	-68.2	-38.4	-177.2	V
404	GLY	HHH H	401,408		0.8	0.01	47.3	0.75	-62.3	-40.1	-178.9	V
406	GLN	HHH H	402,409		1.6	0.01	179.8	1.29	-60.7	-40.4	-179.2	V
411	ILE	HHH H	407,415		9.8	0.05	105.4	0.65	-63.4	-38.7	-175.9	V
414	PHE	HHH H	411		0.0	0.00	119.4	0.61	-67.5	-57.3	-177.7	V
415	ALA	CCH C	411,412		42.3	0.34	75.8	0.87	92.5	-87.9	151.9	P,O
416	CYS	BCB B			60.9	0.42	89.4	0.86	-122.3	74.2	149.4	O
417	PHE	BCB B	188,189		115.8	0.51	147.1	0.75	160.5	145.6	-146.6	V,O
418	VAL	BCB B	234,281		1.9	0.01	84.9	0.62	-35.3	164.0	169.0	V,O
419	GLN	CCC C	162		0.0	0.00	83.1	0.60	85.8	176.9	-178.6	V,P
420	ASN	CCC C			0.1	0.00	86.2	0.74	-132.1	56.9	-155.5	V,O
421	LEU	CBC C	233		9.6	0.05	149.7	0.92	-110.7	-167.6	-168.7	O
428	GLU	CBC C			62.6	0.33	178.7	1.34	-101.6	124.7	-178.5	V
431	LEU	CCC C			230.6	1.11	139.2	0.85	-106.5	360.0	360.0	A

**Figura 12: Evaluación del modelo tridimensional de la lipasa 2-17 posterior al proceso de refinamiento, obtenida mediante la herramienta VADAR.** Los errores que se pueden identificar con el servidor utilizado son: **A:** Indica posible problema con ASA4 fraccional (fASA > 1.0). **V:** Indica posible problema con volumen fraccional (fV < 0.8 o fV > 1.2). **P:** Indica posible problema con ángulos Phi y Psi (Morris y col., 1992). **O:** Indica posible problema con ángulo omega (omega > 170 u omega < -170) **C:** Indica enlace peptídico cis (-20 < omega < 20).

El modelo final entregado por Modeller con los 10 *loops* refinados arrojó un total de 158 aminoácidos con problemas en la estructura o con mala evaluación. En comparación con la estructura de la lipasa 2-17 sin refinar, se mejoraron 36 aminoácidos los cuales no presentaron problemas, se mantuvieron 56 aminoácidos con los mismos problemas y aparecieron 102 nuevos aminoácidos con problemas principalmente en el volumen fraccional y con el ángulo omega. Dentro de los aminoácidos mal evaluados, un 9,2% corresponde a aminoácidos que se encuentran estructuras secundarias de hojas β, un 28,1 % corresponde a α hélices y un 62,7% a *loops*.

Debido a que el modelo fue realizado en base a una estructura que presentó un porcentaje de identidad en el límite para poder realizar modelamientos, en el caso de los valores obtenidos para los aminoácidos con problemas en el volumen fraccional y los aminoácidos que presentaron problemas con sus ángulos omega, se descartaron los aminoácidos que se encontraban levemente bajo o sobre los límites aceptables (considerando un porcentaje de error del 10%), y se analizaron sólo los aminoácidos que presentaban problemas por sobre y bajo este porcentaje.

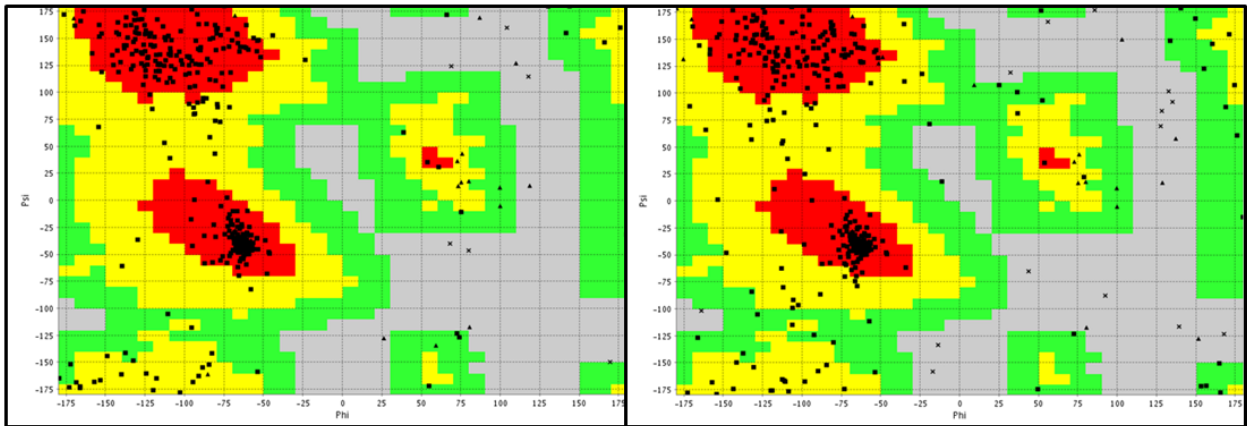
Al analizar los aminoácidos con problemas pero considerando los nuevos límites para los valores definitorios tanto del volumen fraccional como del ángulo omega, en ambos casos (el modelo de la lipasa 2-17 sin refinar y el modelo refinado), se observó una disminución prominente de aminoácidos problemáticos. Para el caso de la lipasa sin refinar, el nuevo número de aminoácidos con problemas descendió a 32, que corresponde al 34,8 % del total de aminoácidos considerados inicialmente. Así también, se observó un notorio cambio para el caso del modelo refinado, en el cual el número de aminoácidos problemáticos descendió a 77, es decir un 48,7% de los aminoácidos seleccionados en primera instancia.

Esto nos indica que si bien el modelo presenta bastantes aminoácidos con problemas, más del 50% de estos aminoácidos, para ambos casos, no tienen graves problemas en su ordenamiento espacial y cuyos valores tanto para el volumen fraccional como para el ángulo omega no supera el 10% del valor aceptado. Este es un buen indicador de la calidad del modelo, asumiendo que se evaluó la estructura proteínica completa.

Para solucionar el problema presente en los ángulos omega de algunos de los aminoácidos utilizando el software Swiss-pdb viewer (Guex & Peitsch, 1997; Kaplan & Littlejohn, 2001) se forzó la molécula para que estos tuviesen sus ángulos dentro del rango. Sin embargo, este cambio provocó un deterioro en la calidad de la estructura, apareciendo problemas en otros residuos, por lo que se descartó forzar el ángulo omega a estar dentro del rango de estos aminoácidos (datos no mostrados).

#### **2.4.8 Comparación de los Gráficos de Ramachandran para el modelo tridimensional de la lipasa 2-17 antes y después del refinamiento.**

Para estudiar un poco más a fondo la posición de los aminoácidos del modelo tridimensional de la lipasa 2-17, se realizó un gráfico de Ramachandran tanto para el modelo refinado como el modelo sin refinamiento. Este gráfico permite tener una idea de la participación de cada aminoácido en cada estructura. En él se pueden visualizar todas las combinaciones posibles de ángulos dihédricos  $\Psi$  ( $\psi$ ) y  $\Phi$  ( $\phi$ ) de los aminoácidos de la lipasa y es posible realizar una comparación entre los modelos, refinado y sin refinar, ya que existen combinaciones de ángulos típicas para cada estructura ( $\alpha$ - hélices y hojas  $\beta$ ).



**Figura 13: Gráficos de Ramachandran del modelo tridimensional de la lipasa 2-17- Izquierda:** Modelo sin refinar. **Derecha:** Modelo refinado. En el primer cuadrante se encuentran las combinaciones de las estructuras hélice  $\alpha$  con giro hacia la izquierda; el segundo cuadrante muestra las combinaciones de la hoja  $\beta$  y en el tercer cuadrante las de las estructuras  $\alpha$  hélices con giro hacia la derecha y los *loops*. Los triángulos, los cuadrados y las cruces, representan a las glicinas, los aminoácidos permitidos y los aminoácidos no permitidos, respectivamente.

Al observar ambos gráficos es posible determinar que en el modelo refinado existe un aumento de aminoácidos que no se encuentran atribuidos a ninguna estructura secundaria, ya sea  $\alpha$  hélice o bien hoja  $\beta$ , posiblemente debido a que el refinamiento realizado por Modeller para mejorar el posicionamiento de los *loops*, afectó otros aminoácidos, fundamentalmente ubicados en los extremos de las estructuras  $\alpha$  hélices (Figura 13).

Al analizar y comparar el gráfico de Ramachandran del modelo tridimensional de la lipasa 2-17 una vez realizado el refinamiento, con el obtenido a partir del modelo sin refinar, es posible establecer el que el número de aminoácidos que no tienen una estructura secundaria atribuida aumenta una vez refinado el modelo .

Finalmente, se realizaron las evaluaciones de cada uno de modelos con *loops* refinados de manera particular y fueron comparados con la evaluación de la proteína sin refinar. Para cada uno de los casos, el mejor evaluado resultó ser el modelo de la lipasa previo al refinamiento, indicando nuevamente la intervención sobre aminoácidos presentes en estructuras secundarias. Esto es debido a que tanto las estructuras  $\alpha$  hélices como la hoja  $\beta$  central, que fueron modeladas en base a una estructura obtenida por cristalografía de rayos X, tenían efectivamente una correcta orientación en comparación a las posibles orientaciones de los aminoácidos en los *loops*. Es por ello, que programas como VADAR que realizan una evaluación global y local de los aminoácidos en la proteína, penalizan fuertemente los movimientos de estos aminoácidos ubicados en estructuras tipo  $\alpha$  hélices u hojas  $\beta$ , en comparación a la valorización entregada por el correcto posicionamiento de los aminoácidos en un *loop* al hacer el refinamiento, lo que conlleva a una mala evaluación final del modelo.

#### 2.4.9 Validación del modelo tridimensional de la Lipasa 2-17 utilizando el servidor MolProbity

Finalmente, utilizando el servidor MolProbity, se procedió a validar el modelo obtenido con Modeller. MolProbity es un servidor cuyo objetivo general es validar la calidad de estructuras tridimensionales de proteínas, ácidos nucleicos y complejos. Proporciona tanto un análisis detallado de los contactos de los átomos con problemas estéricos dentro de las moléculas, así como el diagnóstico de los ángulos dihedrales. Además calcula y muestra los enlaces de hidrógenos y las fuerzas van der Waals en las interfaces entre los componentes (Chen y col., 2009).

Nuevamente se llevó a cabo un análisis sobre los ángulos phi y psi permitidos, generando el gráfico de Ramachandran (datos no mostrados), para determinar el valor phi y psi de los residuos y a partir de éstos establecer si los residuos se ubican en zonas con o sin impedimento estérico.

El gráfico de Ramachandran permite aproximar la estructura tridimensional de la proteína, a partir de las combinaciones phi-psi presentes. En el caso de la lipasa 2-17 la mayoría de sus residuos presentan valores  $-50 < \phi < -180$  y  $70 < \psi < 180$  (cuadrante superior izquierdo de la gráfica) y  $-30 < \phi < -150$  y  $-10 < \psi < -60$  (cuadrante inferior izquierdo de la gráfica), indicando una predominancia de residuos en estructuras secundarias tipo hoja  $\beta$  y  $\alpha$  hélices, lo cual es concordante con antecedentes bibliográficos para este tipo de estructuras (Jung y col., 2008; Kim y col., 1997). Los aminoácidos más afectados fueron Lys21, Leu48, Arg94, Ser125, Gln250, Lys272, Ala289, Ser331 y Phe417, todos aminoácidos ubicados en *loops* de la proteína.

La Tabla II muestra el resumen del análisis de calidad para la estructura final de la lipasa 2-17. Existe un 6,35% de rotámetros pobres, es decir, aquellos residuos que poseen problemas en la rotación de sus cadenas laterales sobre el esqueleto rígido polipeptídico. Por otro lado un 0,65% de residuos presentan problemas en sus ángulos.

**Tabla II:** Tabla resumen del análisis de la calidad de la estructura realizado con MolProbity.

Análisis de la calidad de la estructura realizado por MolProbity		
Parámetros	Lipasa 2-17	Objetivo
Rotámetros libres	6.35%	<1%
Ramachandran no permitidos	3.03%	<0.05%
Ramachandran favorecidos	89.04%	>98%
Desviación $C\beta > 0.25\text{Å}$	2.22%	0
Enlaces incorrectos en cadena principal	0.00%	0%
Ángulos incorrectos en cadena principal	0.65%	<0.1%

Al analizar cada uno de los parámetros es posible determinar que para los rotámetros pobres, sólo 7 de 23 problemas corresponden a aminoácidos presentes en  $\alpha$  hélices, mientras que para los problemas

presentados en *ramachandran outliers* dos correspondieron a aminoácidos presentes en hélices  $\alpha$  y dos a hojas  $\beta$ , de los 13 problemas detectados por Molprobit. Finalmente para las desviaciones  $C\beta$  y los ángulos incorrectos en la cadena principal un bajo porcentaje de los aminoácidos problemáticos se encuentran en  $\alpha$  hélices y hojas  $\beta$ .

Si bien la mayoría de los problemas están presentes en los *loops* de la proteína (problema habitual en los modelos por comparación), pudiendo afectar el posicionamiento y espaciado de las hojas  $\beta$  y de las  $\alpha$  hélices, al hacer el análisis de cada uno de los problemas, no se vio afectada la formación de la estructura secundaria, ya que los aminoácidos seleccionados no se encuentran en los *loops* principales. Nuevamente estos problemas reflejan más bien la falta de identidad entre los diferentes modelos y la carencia de un molde para modelar ciertas áreas de la proteína.

#### **2.4.10 Modelamiento a través de Swiss Model**

Debido a las dificultades que se presentaron al modelar la estructura tridimensional de la lipasa 2-17, principalmente por la poca información en las bases de datos sobre estructuras tridimensionales de lipasas psicrófilas obtenidas por cristalografía de rayos X o RMN, o muy baja identidad en las secuencias, el servidor Swiss Model, que integra bases de datos y construye un modelo tridimensional de la proteína, es una buena herramienta para comparar el modelo tridimensional de la lipasa 2-17 obtenido mediante Modeller (Biasini y col., 2014; Kiefer y col., 2009; Guex y col., 2009).

Al igual que Modeller, la construcción exitosa del modelo creado por Swiss Model, requiere de al menos una estructura tridimensional determinada experimentalmente que presente una secuencia aminoacídica significativamente similar a la proteína blanco. Así, el modelamiento de Swiss Model comprende las mismas cuatro etapas que Modeller y cada una es repetida hasta satisfacer los resultados del modelamiento.

- (i) La identificación de los moldes estructurales
- (ii) Un alineamiento de la secuencia blanco y las estructuras de los moldes
- (iii) La construcción del modelo
- (iv) La evaluación de la calidad del modelo.

Para comenzar el modelamiento se ingresó la secuencia aminoacídica de la Lipasa 2-17 y se procedió a elaborar un primer modelamiento. El servidor Swiss Model entregó dos modelos, pero solo uno de ellos fue utilizado. El primer modelo entregado fue elaborado en base al molde 1jjiB considerando aminoácidos comprendidos entre el 123 y el 418 cuya evaluación resultó ser positiva, mientras que el modelo 2, se elaboró a partir del molde cuyo PDB ID es 1lzlA considerando los

aminoácidos desde el 62 hasta el 424, pero su evaluación resultó ser negativa y poco recomendado, por lo tanto se realizó el trabajo comparativo con el modelo 1 (Tabla III).

Swiss Model utiliza diferentes herramientas para realizar una completa evaluación del modelo, entregando una estimación tanto global como local de la calidad del modelo (Schwede y col., 2003).

Para estimar la calidad global del modelo, se utiliza la herramienta de puntuación global QMEAN4. Esta función de puntuación es una combinación lineal de cuatro descriptores estructurales utilizando potenciales estadísticos: La geometría local es analizada a través del ángulo de torsión potencial sobre tres aminoácidos consecutivos. Dos potenciales de interacción dependientes de la distancia se utilizan para evaluar las interacciones de largo alcance: la primera es una aplicación a nivel de los residuos basada solo en los átomos C-beta y la segunda un potencial de todos los átomos, el cual es capaz de capturar más detalles del modelo. Además, es posible analizar la ubicación espacial y profundidad de los átomos en la proteína en base a un potencial de solvatación. QMEAN4 es una puntuación de fiabilidad del modelo completo que se puede utilizar con el fin de comparar y clasificar modelos alternativos elaborados en base a un mismo modelo dado (Arnold, y col., 2006).

Adicionalmente, Swiss Model entrega un puntaje de los componentes individuales de QMEAN y una medida de la calidad absoluta del modelo entregando un estimativo del grado de naturalidad de las características estructurales observadas en un modelo y describiendo la probabilidad de que un determinado modelo sea de calidad comparable a las estructuras experimentales (Benkert y col. 2008).

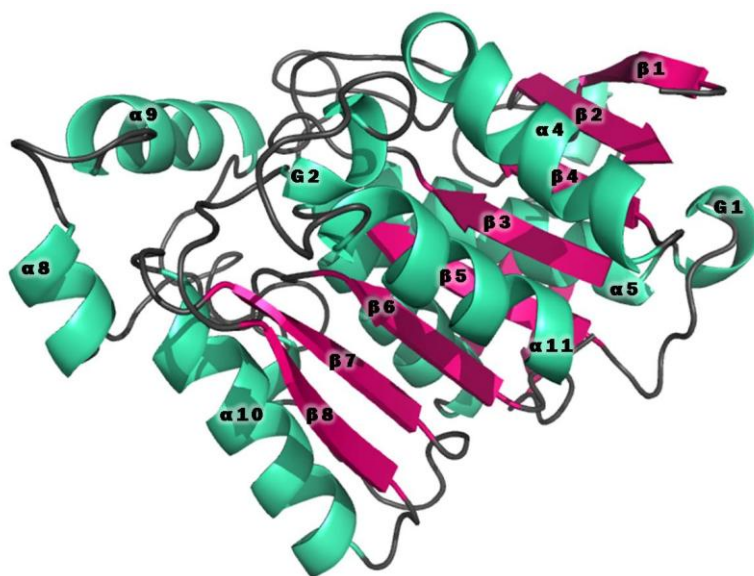
Por último, este programa entrega puntajes locales, cuyos residuos están coloreando por error y una gráfica de error los residuos analizados.

Para estimar la calidad de los modelos locales, Swiss Model evalúa la calidad del empaquetamiento del modelo a través la herramienta ANOLEA (*atomic empirical mean force potential*). El programa realiza los cálculos de energía de una cadena proteica, evaluando el medio ambiente no local de cada átomo pesado en la molécula.

**Tabla III: Información del modelo obtenido por Swiss Model.**

<b>Identificación del modelo obtenido por Swiss Model</b>	
<b>Información del Modelo</b>	<b>Lipasa 2-17</b>
Rango de residuos modelados	123 a 418
Molde utilizado	1jjiB (2.20 Å)
Identidad de Secuencia [%]	31,31%
<i>E-value</i>	0,00E+00
Puntaje QMEAN Z	-3,08





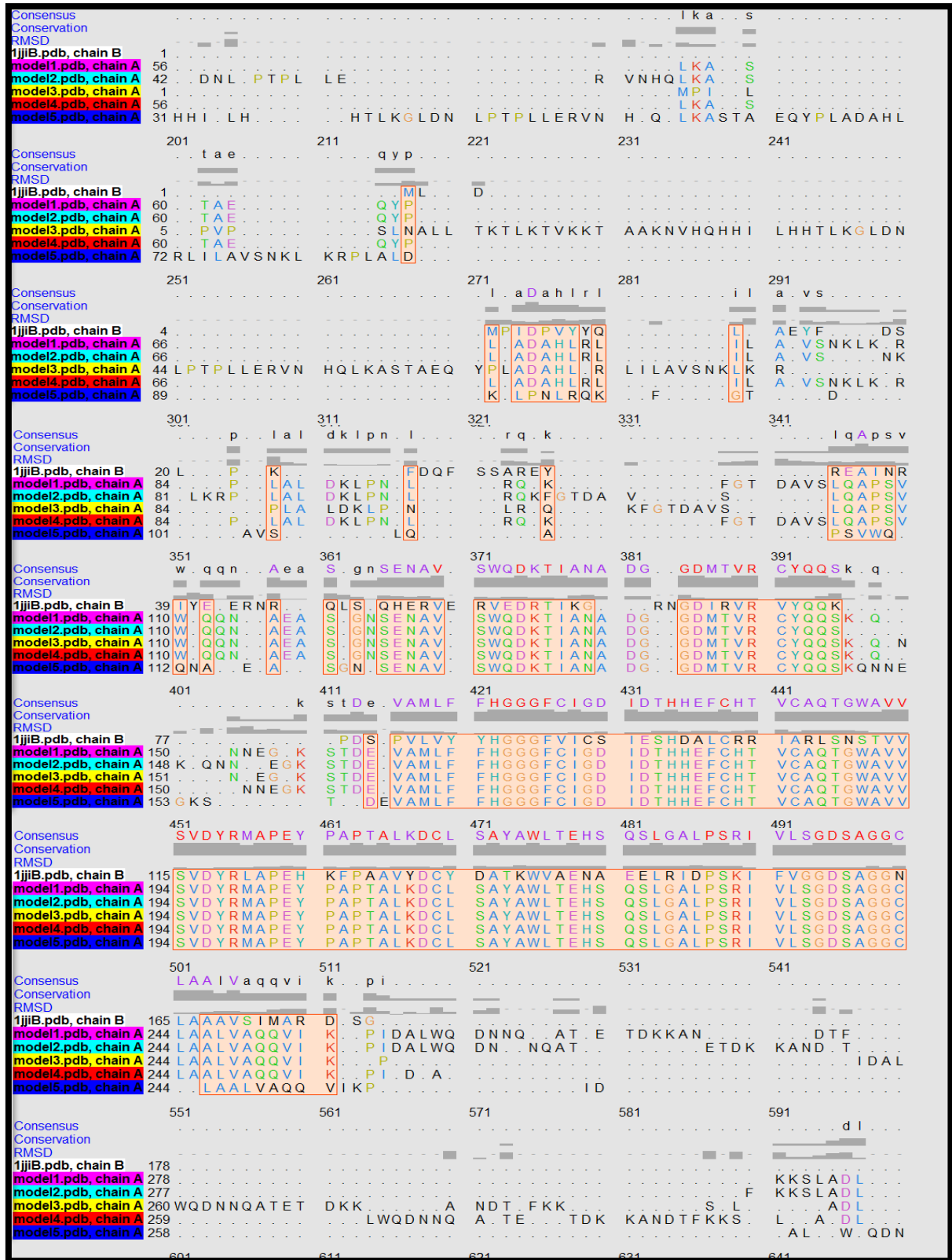
**Figura 14: Modelo tridimensional de la lipasa 2-17 obtenido mediante Swiss Model** Las  $\alpha$  hélices se encuentran de color verde mientras que las hojas  $\beta$  se encuentran de color rosa. En color negro se encuentran indicados los *loops*. Figura realizada con el programa Pymol.

A pesar de que el modelo final obtenido por Swiss Model (Figura 14) está compuesto por 293 aminoácidos (desde Ala123 a Val416), es posible observar una estructura secundaria central compuesta por 8 hojas  $\beta$ , rodeada por 10 estructuras  $\alpha$  hélices formando una estructura tipo barril. A diferencia del modelo obtenido por Modeller, Swiss Model estableció dos hebras  $\beta$  que se encuentran fuera de la estructura tipo barril y que posiblemente no forman parte de una posible tapa en el fenómeno de activación interfacial, debido a que se encuentran alejadas de la estructura central de la lipasa donde debiese estar el sitio activo.

#### 2.4.11 Modelamiento a través de I-TASSER

Finalmente para comparar el modelo tridimensional de la lipasa 2-17 obtenido por Modeller, se realizó un modelamiento mediante la herramienta computacional I-TASSER que también permite la predicción de la función y estructura de proteínas (Roy y col., 2010). La construcción de los modelos tridimensionales está basada en alineamientos múltiples por LOMETS (*local meta-threading server*) y simulaciones tipo TASSER (*threading assembly refinement*), haciendo coincidir los modelos predichos con bases de datos de función de proteínas (Roy y col., 2012; Zhang, 2008).

Una vez realizado el modelamiento, el servidor I-TASSER arrojó cinco modelos. Éstos fueron analizados y reevaluados. Para seleccionar uno de los cinco modelos, lo primero fue determinar qué estructuras secundarias en los 5 modelos eran comunes en cada uno de ellos, y se realizó un alineamiento de los modelos según su estructura (Figura 15).



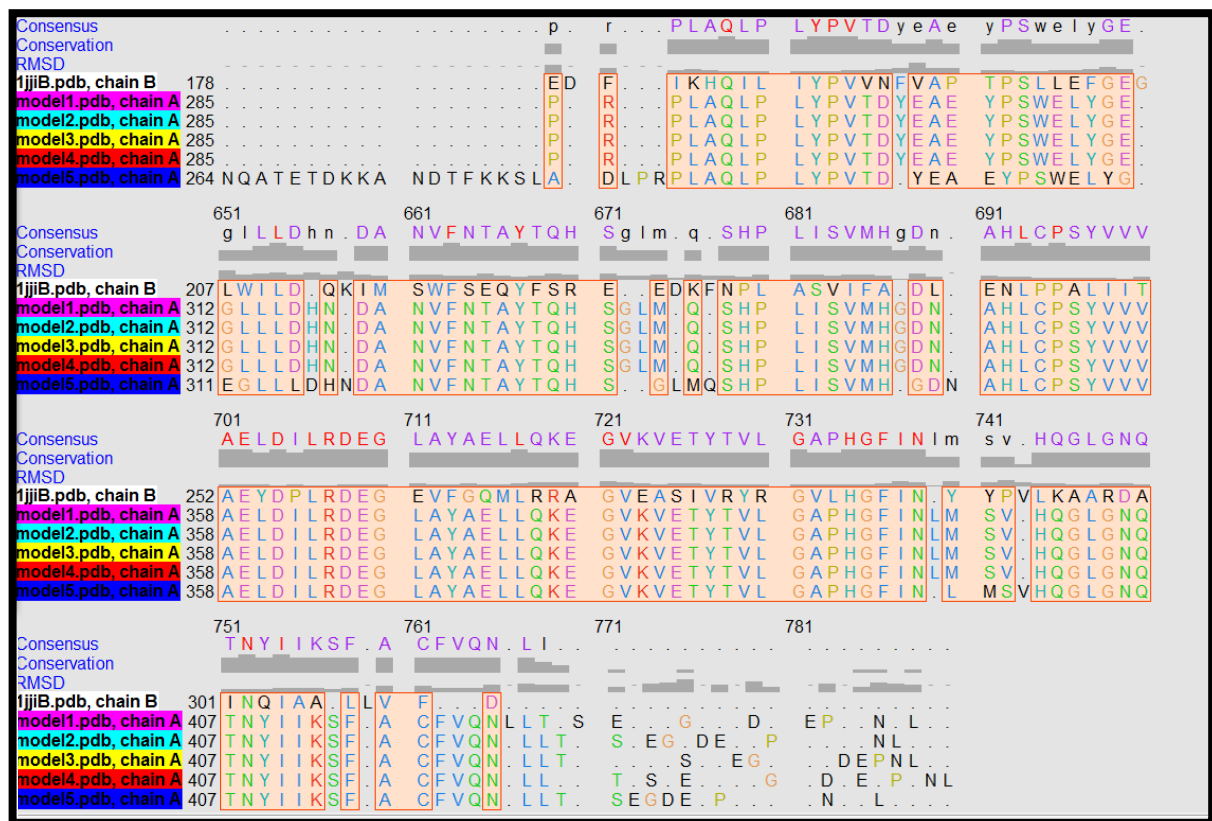
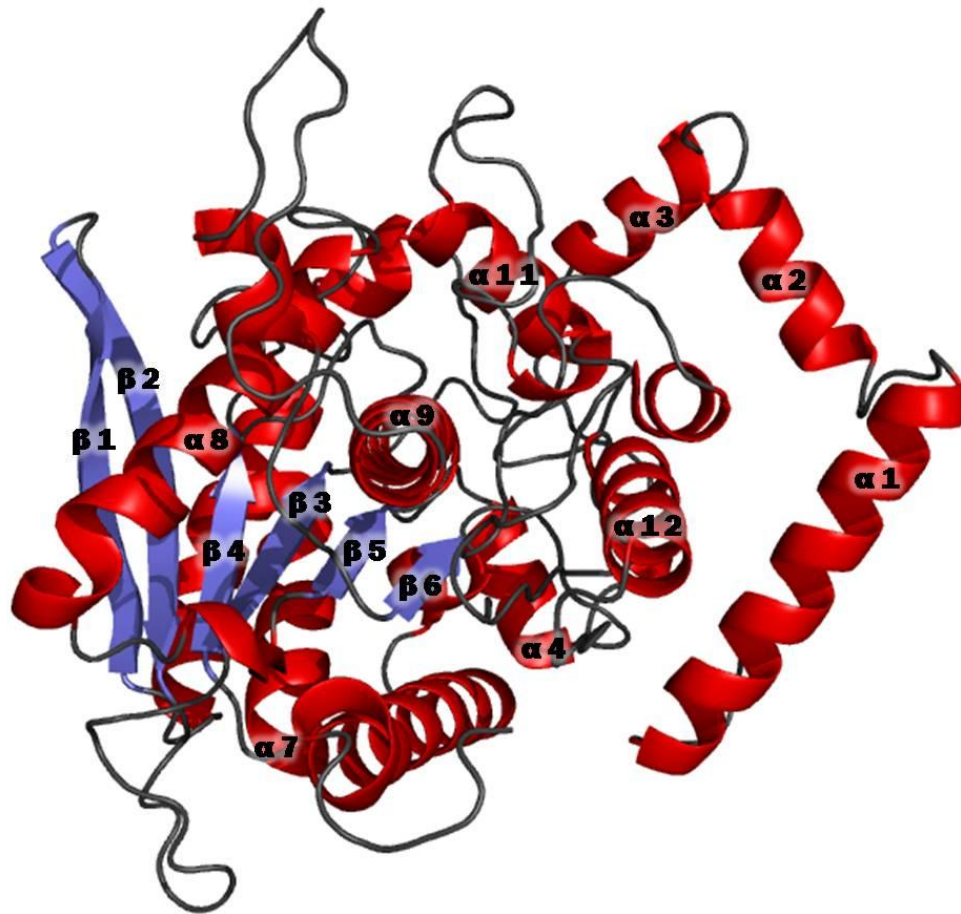


Figura 15: Alineamiento de estructuras secundaria de los cinco modelos tridimensionales de la lipasa 2-17 obtenido por I-TASSER. En anaranjado se identifican los aminoácidos que se encuentran igualmente posicionados en cada uno de los modelos. Alineamiento realizado con el programa Chimera.

2.4.11.1 Evaluación y selección de los modelos obtenidos por I-TASSER

Usando las herramientas Verify 3D, Procheck y MolProbity, se realizó una evaluación de cada uno de los modelos obteniendo el mejor resultado para el modelo 1 (datos no mostrados). Particularmente, Procheck evalúa la calidad estereoquímica de la estructura de una proteína, entregando gráficas que analizan residuo por residuo y la geometría general del modelo. Por otro lado, Verify3D analiza la compatibilidad de un modelo atómico (3D) con su propia secuencia aminoacídica (1D). A cada residuo se le asigna una clase estructural en función de su ubicación y su entorno (alfa, beta, loop, polar, no polar, etc.) y una serie de estructuras correctamente evaluadas se utilizan como referencia para obtener una puntaje para cada uno de los 20 aminoácidos en esta clase estructural (Bowie y col., 1991; Luethy y col., 1992).

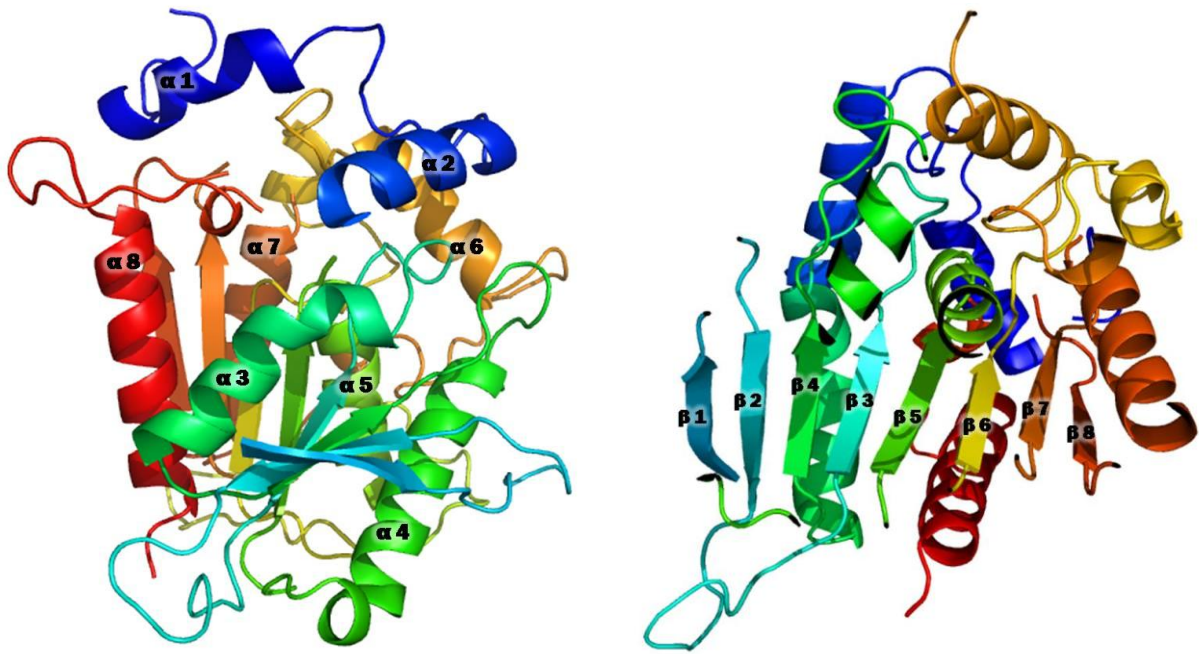


**Figura 16: Modelo tridimensional de la lipasa 2-17 obtenido mediante I-TASSER.** Las  $\alpha$  hélices se encuentran de color rojo mientras que las hojas  $\beta$  se encuentran de color azul. En color negro se encuentran indicados los *loops*. Figura realizada con el programa Pymol.

Como era de esperar, al igual que en los otros modelos, la zona central del modelo, está formado por 8 hojas  $\beta$  y las  $\alpha$  hélices se encuentran rodeando a estas hojas, formando una estructura tipo barril (Figura 16).

#### 2.4.12 Modelo final de la lipasa 2-17

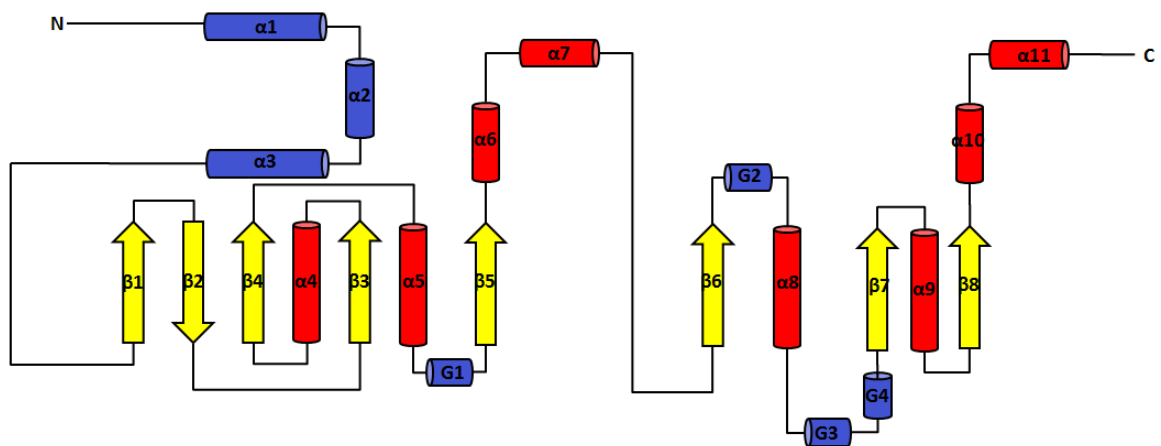
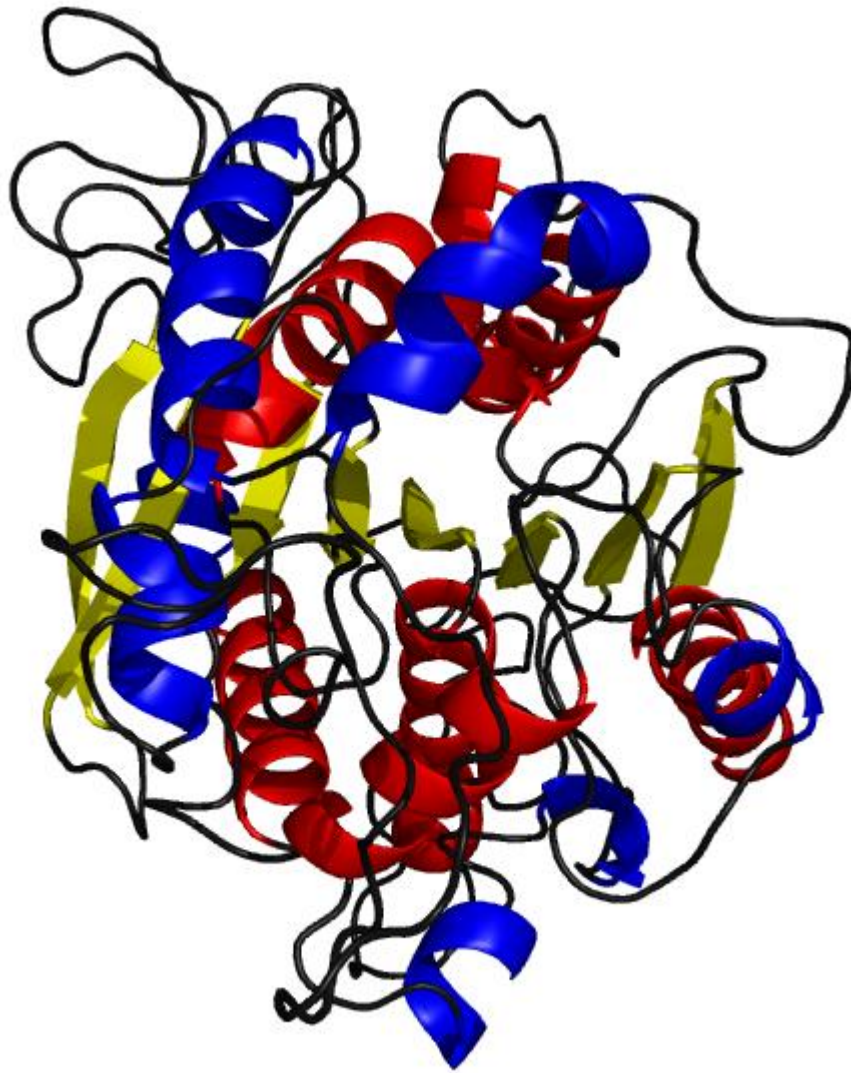
Como se predijo gracias a las distintas herramientas de modelamiento y fue respaldado con la gráfica de Ramachandran, el modelo obtenido para la lipasa 2-17 presenta estructura secundaria, formada principalmente por hojas  $\beta$  y  $\alpha$  hélices. En todos los modelos es posible observar que la estructura tridimensional de la lipasa 2-17 presenta en el centro 8 hojas  $\beta$  rodeadas por  $\alpha$  hélices. El modelo final seleccionado, correspondió al elaborado con el programa Modeller previo al proceso de refinamiento, ya que este modelo mostró los mejores resultados entregados por las diferentes herramientas evaluativas utilizadas (VADAR, Procheck y Verify 3D).



**Figura 17: Modelo tridimensional final de la lipasa 2-17 obtenido a través de Modeller. Izquierda:** Modelo con las principales  $\alpha$  hélices indicadas. **Derecha:** Modelo con las principales hojas  $\beta$  indicadas. Ambas figuras están coloreadas de acuerdo a la posición en la secuencia, desde el N-terminal en azul hasta el C-terminal en rojo. Figura realizada con el programa Pymol.

Particularmente, como es observado en estructuras de  $\alpha/\beta$  hidrolasas, la estructura de la lipasa está conformada por ocho hojas  $\beta$  centrales entre las cuales solo una se encuentra ubicada de manera antiparalela (Figura 17). La hoja  $\beta 1$  se une directamente con la  $\beta 2$  que se encuentra ubicada de forma antiparalela al resto de las hojas que componen la estructura central de la lipasa. La hoja  $\beta 2$  se une a la subsiguiente hoja, correspondiente a la hoja  $\beta 3$ , que restablece la posición original de las hojas  $\beta$ . La hoja  $\beta 3$  se une a una  $\alpha$  hélice que continúa en la hoja  $\beta 4$ , que se encuentra entre las hojas  $\beta 2$  y  $\beta 3$ . A partir de la Hoja  $\beta 4$  la conexión entre las siguientes hojas se ve interrumpida por una  $\alpha$  hélice en cada uno de los casos hasta llegar a la hoja  $\beta 8$ .

En la Figura 18, se muestra el modelo de la estructura tridimensional lipasa 2-17, obtenido por comparación, utilizando como molde la estructura de la lipasa AFEST (PDB ID 1JJI). El dominio central de la lipasa 2-17 está formado desde los residuos W126 hasta A415 y la conformación de la enzima correspondió a una conformación abierta, es decir, con el sitio activo accesible para el sustrato.



**Figura 18: Plegamiento de la Lipasa 2-17.** Arriba: Plegamiento característico de Lipasa 2-17. Abajo: Representación esquemática del plegamiento de la Lipasa 2-17. Las hojas  $\beta$  se encuentran representadas por flechas amarillas, las  $\alpha$  hélices centrales por cilindros rojos y las  $\alpha$  hélices superficiales y  $\beta_{10}$  por cilindros azules.

La estructura de hoja  $\beta$  central se encuentra rodeada de seis estructuras hélices  $\alpha$ ,  $\alpha 4$  (I174-T188),  $\alpha 5$  (T207-E221),  $\alpha 6$  (A240-A249),  $\alpha 8$  (A317-G327),  $\alpha 9$  (L363-E377) y  $\alpha 11$  (Q401-A415), que corresponden a las estructuras  $\alpha$  hélices  $\alpha A$ ,  $\alpha B$ ,  $\alpha C$ ,  $\alpha D$ ,  $\alpha E$ ,  $\alpha F$  respectivamente. La principal diferencia en comparación al clásico plegamiento  $\alpha/\beta$  hidrolasa es la presencia de dos hélices  $\alpha$  entre las  $\beta 5$  y  $\beta 6$  (Figura 18). Además la lipasa 2-17 se encuentra formada por ocho estructuras  $\beta$ :  $\beta 1$ ,  $\beta 2$ ,  $\beta 3$ ,  $\beta 4$ ,  $\beta 5$ ,  $\beta 6$ ,  $\beta 7$  y  $\beta 8$ , once estructuras hélices  $\alpha$ , y cuatro hélices  $3_{10}$  denominadas con la letra G (Tabla IV).

**Tabla IV: Estructuras secundarias de la lipasa 2-17 y sus aminoácidos correspondientes**

Correspondencia de estructura secundaria y aminoácidos que la componen	
Estructura Secundaria	Aminoácidos
$\alpha 1$	9-20
$\alpha 2$	61-67
$\alpha 3$	75-78
$\beta 1^*$	125-132
$\beta 2^*$	138-145
$\beta 3^*$	160-164
$\alpha 4^*$	174-188
$\beta 4^*$	191-194
$\alpha 5^*$	207-221
G1*	223-226
$\beta 5^*$	232-237
$\alpha 6^*$	240-249
$\alpha 7$	254-259
$\beta 6^*$	290-293
G2*	304-308
$\alpha 8^*$	317-327
G3*	332-336
G4*	342-344
$\beta 7^*$	353-357
$\alpha 9^*$	363-377
$\beta 8^*$	381-385
$\alpha 10$	393-399
$\alpha 11^*$	401-415

\* Estructuras secundarias del dominio central de la lipasa 2-17.

Por otra parte, en la lipasa 2-17, la zona de contacto lipídica contiene varios residuos hidrofóbicos grandes como Phe, Ile y Leu. Se ha estudiado que algunos de los residuos Phe e Ile en otras lipasas presentan una función de anclaje de lípidos, y por lo tanto probablemente penetran en la parte

hidrofóbica de la superficie de los lípidos, pudiendo reflejar en este caso, una posible estructura tipo tapa. La tapa es también una parte central de la zona de contacto lipídica de la mayoría de las triacilglicerol lipasas. El desplazamiento de la tapa aumenta el área de superficie hidrofóbica de las lipasas dentro del área de la zona de contacto lipídica. Las estructuras secundarias de la lipasa 2-17 que podrían actuar como tapa y participar en el fenómeno de activación interfacial se encuentran descritas en el capítulo 5.



## 2.5 CONCLUSIONES

El modelamiento por homología es una herramienta que permite construir modelos útiles de la estructura tridimensional de una proteína cuya estructura es desconocida. Para construir este modelo se necesita conocer la secuencia aminoacídica de la proteína y la estructura tridimensional de alguna proteína homóloga, evolutivamente relacionada. En primera instancia, para obtener el modelo de la estructura de la lipasa estudiada se utilizó una lipasa cuyo porcentaje de identidad era de aproximadamente un 40%. Si bien este es un valor límite para construir modelos, es útil a la hora de armar un modelo general. Se debe considerar al momento de realizar el modelo, la importancia que tiene un buen sistema de evaluación para el modelo, ya que esto permite determinar qué aminoácidos deben ser re-posicionados, qué regiones deben ser refinadas y qué estructuras deben ser mejoradas. Debido a que existe una posible fuente de errores al modelar una enzima, la evaluación del modelo final es fundamental para establecer qué fase se debe repetir y mejorar, en caso de ser necesario.

Particularmente el modelo obtenido para la lipasa en estudio presentó una evaluación positiva lo que permitió trabajar con él para caracterizar posteriormente a esta enzima. A pesar que se utilizó una proteína molde termófila, las bases estructurales globales de las diferentes lipasas son similares en términos de estructuras secundarias y terciarias (Gianese y col., 2001). Básicamente, las lipasas están formadas por una estructura  $\alpha/\beta$  hidrolasa conformada por ocho hebras  $\beta$  centrales que forman una estructura semicilíndrica, entre las cuales siete se encuentran ubicadas de forma paralela, mientras que solo una de manera antiparalela. Con el modelo obtenido para la lipasa 2-17 fue posible apreciar la forma de barril que permite que el sustrato entre al sitio activo para ser hidrolizado. Al mirar la estructura de la lipasa estudiada, es posible apreciar que en la zona central, hay una hendidura hidrofóbica que podría participar en la catálisis enzimática. Esta hendidura podría corresponder al bolsillo hidrofóbico que permite el encaje de las cadenas de los ácidos grasos que son cortadas al hidrolizar el sustrato. Además la lipasa 2-17, cuenta con un par de  $\alpha$  hélices que podrían corresponder a la tapa de la enzima, necesaria para la activación interfacial característica de este tipo de enzimas, que depende de la concentración micelar del solvente. En la tapa, el área hidrofóbica se pone en contacto con el solvente estabilizando la estructura y permitiendo la entrada del sustrato al sitio activo de la enzima.

Todos los modelos obtenidos con las diferentes herramientas de modelamiento por homología (Modeller, Swiss Model e I-TASSER) exhibieron estructuras básicamente idénticas, cuyas principales diferencias radican en las zonas de los *loops*, en estas zonas el porcentaje de identidad con la proteína molde fue básicamente cero.

## **CAPÍTULO 3**

**Identificación de la Triada Catalítica de la Lipasa 2-17**

### 3.1 RESUMEN

Las lipasas son éster-hidrolasas carboxílicas que rompen los enlaces éster de los acilglicéridos mediante la adición de una molécula de agua, dando lugar a ácidos grasos libres y glicerol. El centro activo de estas enzimas tiene tres aminoácidos catalíticos cuya posición dentro del plegamiento suele ser muy conservada: una serina nucleofílica, un ácido aspártico y una histidina. La serina catalítica está generalmente incluida en el pentapéptido conservado Gly(Ala)-Xaa-Ser-Xaa-Gly, el cual forma un giro entre una hoja  $\beta$  y la siguiente hélice  $\alpha$  cuya función es estabilizar y orientar la serina nucleofílica. La interacción de estos aminoácidos con el sustrato muchas veces es estudiada mediante *docking* molecular, herramienta computacional bastante útil para identificar moléculas que son complementarias a los sitios activos de las enzimas. Particularmente, este sistema de acoplamiento ha sido desarrollado para predecir la posible posición y orientación de estos ligandos en el sitio activo de cada proteína. Para ello, se aplica una función de puntajes para la evaluación de las correctas interacciones en todos los posibles estados. Esta función de puntajes es derivada de mecanismos moleculares, que requieren de unos cuantos parámetros para entregar rápidos resultados de acoplamiento.

En este capítulo, para identificar residuos potencialmente involucrados en la actividad lipolítica, mediante un alineamiento de secuencia con la proteína molde utilizada para el modelamiento, se determinaron tres candidatos que podrían formar la triada catalítica: Serina 239, Aspartato 361 e Histidina 391. Cada uno de estos aminoácidos fue analizado y evaluado estructuralmente. Los análisis computacionales mostraron la cercanía espacial entre ellos, y la posibilidad de establecer interacciones para llevar a cabo la hidrólisis del sustrato. Para estudiar la interacción de cada uno de estos aminoácidos con el sustrato, se realizaron simulaciones computacionales para determinar los sitios de unión de cuatro tipos diferentes de sustratos: triacilglicéridos compuestos por cadenas de 6, 8, 10 y 16 carbonos. Los resultados identificaron al sustrato compuesto por ácidos grasos de 6 carbonos como el mejor y más preferido por la enzima, en comparación a los otros tres cuyas energías de interacción fueron notoriamente menores.

Finalmente se estudiaron 5 mutaciones que podrían explicar la importancia catalítica de estos aminoácidos y para conocer el comportamiento de los variantes de la lipasa nativa, se simuló mediante *docking* molecular la interacción de cada una de las enzimas mutantes con cada uno de los sustratos. El objetivo fue básicamente estudiar la promiscuidad de la triada y evaluar si los cambios estructurales afectan a la selectividad de sustratos. Los resultados demostraron que para cada uno de los mutantes, tanto el sustrato de preferencia como los otros tres sustratos, no mantenían interacciones que permitieran la correcta actividad catalítica de la enzima.

## 3.2 INTRODUCCIÓN

La información bioquímica que se ha acumulado a través de los años en conjunto con las técnicas como cristalografía de rayos X, que permiten la obtención de modelos tridimensionales, han permitido que se conozca la arquitectura atómica de muchas enzimas y cómo ésta se relaciona con su función. Además, proporcionan información sobre las posiciones atómicas en el espacio tridimensional y las distancias entre los átomos, lo que permite analizar las interacciones que mantienen la estructura en las proteínas y cómo los ligandos interactúan con ellas. Sin embargo, esta idea de que las proteínas son estructuras estáticas está muy alejada de la realidad. En el caso de las enzimas, aun en ausencia de sustrato, las estructuras vibran y fluctúan constantemente, indicando que son estructuras dinámicas. Estudios sobre el dinamismo de las enzimas (Zavodszky y col., 1998) han propuesto que el proceso catalítico es una secuencia de reacciones en que cada una se acompaña de movimiento y desplazamiento de distintas partes de la proteína. El proceso de catálisis está acompañado de muchos arreglos internos en las enzimas, existiendo un balance entre las fuerzas desestabilizantes y estabilizantes a la vez y permitiendo que muchos residuos de la proteína se oculten y otros se expongan al solvente (Gunasekaran & Nussinov, 2007). Herramientas como *docking* molecular permiten conocer estas interacciones particularmente recreando escenarios que si bien no son reflejo de la realidad dinámica de las proteínas, permite tener un punto de partida para comenzar a conocer aspectos como la actividad enzimática.

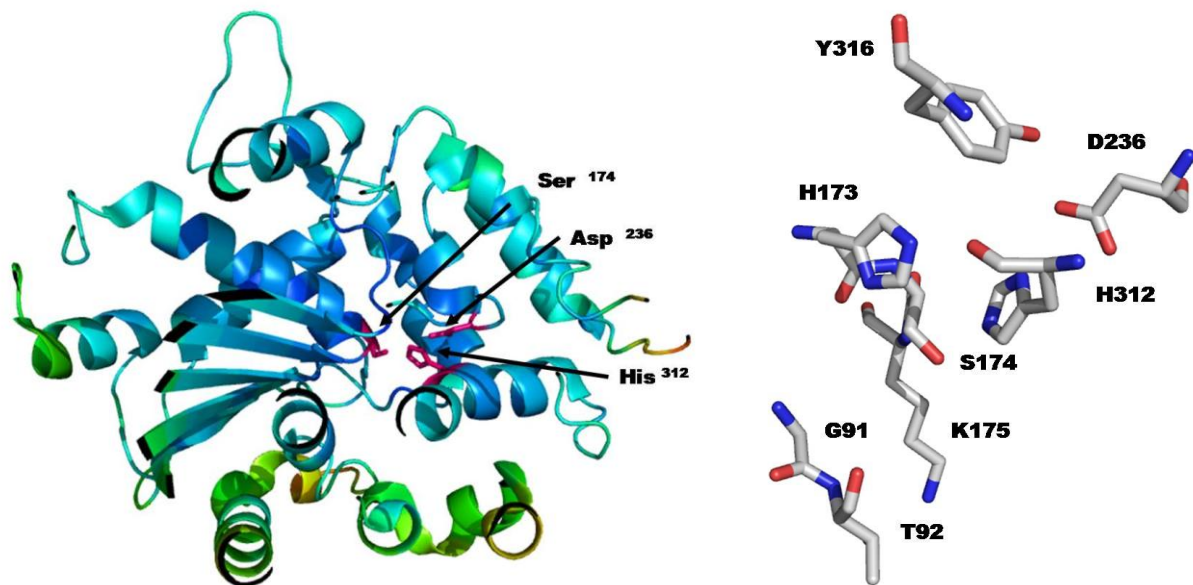
Una enzima activa en solución puede ser representada por una serie de microestados con distintas conformaciones espaciales (Bai y col., 1995), determinadas por diferencias en el plegamiento en puntos específicos de la proteína (Wooll y col., 2000; Mulder y col., 2001). Cada uno de estos microestados tiene afinidades diferentes por los sustratos, incluso en el momento de unión del sustrato al sitio activo, desplazando al equilibrio de los microestados hacia las conformaciones más favorables para llevar a cabo la catálisis (Ma y col., 1999; Kumar y col., 2000). Así, el estado nativo de una proteína se define como la distribución de conformaciones que son ocupadas por la población de moléculas, en donde las variaciones ambientales como temperatura, pH, interacciones moleculares, fuerza iónica, etc pueden intervenir en este estado.

### 3.2.1 Triada catalítica en lipasas

Las enzimas funcionan en base a una amplia variedad de mecanismos. Particularmente, las lipasas catalizan la reacción mediante la hidrólisis de los triacilgliceroles, formando ácidos grasos y glicerol. Una variedad de características estructurales son responsables de la eficiencia catalítica de estas enzimas, pero a pesar de esto, todas las lipasas contienen tres aminoácidos en sus sitios activos: una serina, una

histidina y un aspartato. Estos aminoácidos comprenden la triada catalítica característica de estas enzimas. Comúnmente, estos tres aminoácidos se encuentran ubicados a lo largo de toda la secuencia, pero a nivel estructural se encuentran muy próximos, lo que favorece la hidrólisis del sustrato. Algunas lipasas bacterianas son estructuralmente muy similares a las lipasas de mamíferos, particularmente por la presencia de la serina, fundamentando la hipótesis de que la estructuración de la triada catalítica con la serina se produjo antes de la divergencia evolutiva entre procariontes y eucariontes (Hide W., y col 1992). Por otro lado, al comparar los sitios catalíticos de diferentes lipasas, es posible ver que si bien tienen los mismos residuos catalíticos, la estructura del sitio activo es completamente diferente, lo que sugiere que este mecanismo de las lipasas ha sido establecido en más de una ocasión a nivel evolutivo. Si bien las lipasas difieren en sus secuencias y en sus especificidades de sustratos, la mayoría exhibe la misma triada catalítica.

En lipasas, la serina de su sitio activo está situada en el medio de un pentapéptido de secuencia G-X-S-X-S, mientras que en la mayoría de las esterasas, incluyendo muchas de las lipasas, la serina nucleofílica se encuentra ubicada en el pentapéptido de secuencia G-X-S-X-G. Particularmente en lipasas, esta última secuencia forma un pliegue e importante *loop*, entre una hoja  $\beta$  y una  $\alpha$  hélice (Hilton & Buckley, 1991; Brumlik & Buckley, 1996).



**Figura 19: Estructura lipasa psicrófila M37. Izquierda:** Diagrama de estructuras secundarias de la lipasa M37. En rojo se muestran los aminoácidos de la triada catalítica. **Derecha:** Región del sitio activo lipasa M37 (modificado de Jung y col., 2008). Figura elaborada con el programa Pymol.

En la mayoría de las lipasas, cuando la enzima está en solución acuosa el sitio catalítico se encuentra oculto, pero se hace accesible al contacto con la interfase lipídica. Esto permite que el sustrato pueda acceder al sitio catalítico y la serina catalítica actúe como nucleófilo (Jager y col., 1992).

### 3.2.2 *Docking Molecular*

El conocimiento de las asociaciones moleculares ayuda a la comprensión de una variedad de vías del metabolismo que tienen lugar en una célula (Duhovny D., y col., 2002). Particularmente, el *docking* molecular, es una herramienta muy importante en biología molecular estructural y cuyo principal objetivo es predecir la forma predominante en que un ligando se une a una proteína cuya estructura tridimensional es conocida.

A pesar que las interacciones de van der Waals, hidrofóbicas y electrostáticas afectan en gran medida la afinidades de unión de las moléculas, dar forma al acoplamiento es una condición necesaria pero difícil. Este problema en el proceso de *docking* (o acoplamiento) es complejo debido a varias razones. Por un lado, los modelos de las moléculas no son exactos ya que muchas veces contienen errores experimentales; por otro, las moléculas por lo general se someten a cambios conformacionales durante la asociación, proceso conocido como ajuste inducido. Por lo tanto, los algoritmos usados en el proceso de *docking* molecular deben ser tolerantes a estos problemas, teniendo como consecuencia que el *docking* molecular sea uno de los procesos más difíciles en la bioinformática estructural (Duhovny D., y col., 2002).

La problemática principal de los métodos para predecir en *docking* molecular es encontrar las uniones de baja energía entre una molécula pequeña o un ligando, y el sitio activo de una macromolécula, o receptor cuya estructura es conocida. Resolver el problema de acoplamiento computacional, requiere de una representación exacta de la energía molecular, así como un algoritmo eficiente para encontrar los tipos de uniones potenciales. Aunque muchos algoritmos computacionales han sido descritos para buscar simultáneamente el espacio conformacional en el receptor y uno o varios ligandos tales como el *annealing*, algoritmos evolutivos, y otras búsquedas, están en gran parte enfocados a los estudios de ligandos individuales (Ewing T. y col., 2001; Duhovny D., y col., 2002).

El objetivo de los algoritmos desarrollados para *docking* molecular es detectar una transformación en una de las moléculas, llevándola a un ajuste óptimo con la otra molécula sin causar un problema estérico. Naturalmente, la optimización depende no solo de la forma geométrica de la molécula, sino también de los criterios biológicos, que representan la estabilidad del complejo resultante entre ambas moléculas (Duhovny D., y col., 2002).

### **3.3 MATERIALES Y MÉTODOS**

#### **3.3.1 Determinación de la Triada Catalítica**

##### **3.3.1.1 Alineamiento de la secuencia de la Lipasa 2-17 con el molde**

Mediante el servidor BLAST (*Basic Local Alignment Search Tool*) se procedió a comparar las secuencias aminoacídicas de la lipasa 2-17 con la de la proteína molde con el objetivo de determinar la relación entre los aminoácidos de la triada catalítica. Mediante la información entregada por la base de datos "*Protein Data Bank*" se establecieron los aminoácidos de la triada catalítica de la lipasa molde AFEST PDB ID 1jji (cadena A) y se buscó su correlación con los aminoácidos de la secuencia de la lipasa 2-17.

##### **3.3.1.2 Análisis del modelo tridimensional de la Lipasa 2-17**

Una vez determinados los posibles aminoácidos de la triada catalítica de la lipasa 2-17 PDBViewer se visualizaron tridimensionalmente mediante los programas Pymol, Chimera y se estudiaron sus relaciones espaciales, sus enlaces, sus distancias atómicas, su proximidad en el posible sitio activo y los aminoácidos cercanos que pudiesen estar participando en la catálisis enzimática.

##### **3.3.1.3 Alineamiento estructural del modelo tridimensional de la Lipasa 2-17 con su molde**

Mediante la herramienta "*structure comparison*" del programa Chimera, se realizó un alineamiento estructural de las lipasas 2-17 y 1jjiA para ver la relación espacial de los aminoácidos seleccionados como candidatos de la triada catalítica.

#### **3.3.2 Mutaciones de la Triada Catalítica**

##### **3.3.2.1 Modelamiento de las mutaciones de la triada catalítica de la Lipasa 2-17.**

Utilizando el programa Modeller 9.14, cada una de las mutaciones fueron recreadas en un nuevo modelo tridimensional de la lipasa 2-17, con el objetivo de estudiar la participación de cada uno de los aminoácidos en la triada catalítica. El modelo de cada mutación fue elaborado en base al molde de la lipasa 1jjiA.

### **3.3.3 Docking Molecular**

#### **3.3.3.1 Evaluación de la unión del sustrato a la lipasa 2-17**

Se evaluó la orientación preferente de la lipasa sobre el sustrato, mediante un análisis de *docking* molecular, llevado a cabo con los programas Autodock y Vina. La preparación de los archivos de entrada fue similar para ambos casos, salvo porque en el caso de Autodock se calcularon los mapas de búsqueda.

#### **3.3.3.2 Preparación de ligandos**

Para la preparación de los archivos se utilizó el programa ADT Tools (Hamelryck & Manderick, 2003). En éste programa, se selecciona el archivo pdb correspondiente al receptor, en este caso el modelo de la lipasa 2-17. Una vez abierto el archivo del receptor, se le agregan los hidrógenos polares y posteriormente las cargas parciales. Finalizado, se guarda el archivo en formato pdbqt.

En el caso del ligando, se carga al programa y luego, se selecciona los enlaces en los que se permitirá rotación. Para el caso de los triacilgliceroles, éstos se definen como rígidos, debido a lo cual fue importante chequear que el número de rotámetros posibles se encuentre en cero. Una vez definidos como rígidos, se guardan en formato ligando\_rigido.pdbqt.

#### **3.3.3.3 Selección y ubicación de la caja**

La caja empleada define los límites de búsqueda para la unión del ligando a la proteína. Ésta debe ser suficientemente grande para permitir la completa rotación del ligando dentro de ella. Para el caso del *docking* ciego, la caja se definió centrada en el receptor, mientras que en *docking* forzado se desplazó el centro de la caja con el objetivo de limitar el espacio de búsqueda al deseado.

#### **3.3.3.4 Generación de los mapas de búsqueda (Autodock)**

Si bien a Vina le basta con definir el ligando, receptor, tamaño y ubicación de la caja, en el caso de Autodock se requiere el cálculo de los mapas de búsqueda, lo cual se lleva a cabo mediante la ejecución del comando "Autogrid". Para ello, es necesario la generación de un archivo en formato .gpf (*grid parameter file*), el cual contiene la información asociada al ligando, receptor, tamaño y ubicación de la caja. Al ejecutar este archivo con Autogrid, se generan los mapas requeridos para la siguiente etapa de *docking* molecular.



### **3.3.3.5 Docking molecular**

Para Vina se generó un archivo de configuración que incluye: el ligando a utilizar (triacilglicerol.pdbqt), la lipasa (lipasa217.pdbqt), dimensiones y ubicación de la caja y parámetros de configuración.

Para Autodock se seleccionó el algoritmo genético Lamarckiano y se dejó el resto de los parámetros de simulación por defecto. Se generó un archivo .dpf (*docking parameter file*) el cual fue ejecutado con Autodock.

### **3.3.3.6 Evaluación de la unión dominio-sustrato de la lipasa 2-17 mutada**

La evaluación de unión de las distintas mutaciones (dominios mutados) con la lipasa, se realizó mediante análisis de *docking* molecular, con el software Vina, utilizando para ello el procedimiento definido anteriormente.

## 3.4 RESULTADOS Y DISCUSIÓN

### 3.4.1 Análisis de la estructura central de la lipasa 2-17

#### 3.4.1.1 Plegamiento de la lipasa 2-17

Las lipasas se caracterizan por poseer una estructura tipo  $\alpha/\beta$  hidrolasa. Como se dijo previamente, esta estructura está conformada generalmente por ocho hojas  $\beta$  centrales, formando una estructura semicilíndrica. En general, entre las estructuras  $\beta 5$  y  $\alpha C$  se encuentra ubicado un pentapéptido conservado que contiene a la serina nucleofílica. Así, la estructura general de las triacilglicerol lipasas se puede describir como una estructura tipo barril, cuya serina activa (que pertenece a la triada catalítica de la enzima) se encuentra ubicada en un *loop*, denominado el codo catalítico. Entre las estructuras  $\beta 7$  y  $\alpha E$  se encuentra el aspartato y en algunos casos glutamato, y finalmente entre las estructuras  $\beta 8$  y  $\alpha F$  se encuentra la histidina catalítica. Estos tres aminoácidos se encuentran posicionados en el sitio activo estratégicamente cerca, permitiendo la hidrólisis del sustrato.

Mediante el programa “PDBViewer” y el programa “Pymol” se logró visualizar la estructura central de la lipasa 2-17 con el objetivo de estudiar la similitud con el plegamiento  $\alpha/\beta$  hidrolasa descrito para la mayoría de las lipasas, lo que permitió realizar posteriormente una comparación de los aminoácidos que participan en las diferentes estructuras, ayudando a dilucidar qué aminoácidos podrían estar participando en la catálisis.

El modelo estructural de la lipasa 2-17 presentó una forma muy parecida a la de una proteína globular. Como se predijo, esta estructura consiste en un dominio central similar al observado en las enzimas con plegamiento  $\alpha/\beta$  hidrolasa y posiblemente cuenta con la presencia de una tapa análoga en la encontrada en las lipasas de la familia H.

El núcleo central se compone de los residuos W126 al A415 y contiene una hoja  $\beta$  central compuesta por ocho hebras, en las cuales solo una, la hoja  $\beta 2$  es antiparalela. Seis hélices  $\alpha$  rodean a la hoja  $\beta$  central, siguiendo el mismo patrón que sigue la mayoría de estas enzimas hidrolíticas.

### 3.4.1.2 Aminoácidos involucrados en el mecanismo catalítico de la lipasa 2-17

- **Alineamiento de secuencias**

Una vez establecido que la estructura central de la lipasa 2-17 correspondía efectivamente a un plegamiento  $\alpha/\beta$  hidrolasa, para determinar qué aminoácidos podrían estar participando en la triada catalítica de la lipasa 2-17, se analizaron los sitios activos de varias lipasas. Como se dijo, el sitio activo de las enzimas con este tipo de plegamiento, consiste en tres residuos catalíticos: un residuo nucleofílico, un residuo ácido catalítico, y un residuo de histidina, siguiendo siempre este orden en la secuencia aminoacídica. Particularmente en lipasas, el residuo nucleofílico es siempre una serina, mientras que el residuo catalítico puede variar entre un aspartato o glutamato.

```

LIPASA-2-17      MPILPVPSLNALLTKTLKTVKKTAAKNVHQHHILHHTLKGLDNLPTPLLERVNHQLKAST 60
1JJIB            -----MLDMPIDPVYYQLAEY- 16
                  :      : : * : **

LIPASA-2-17      AEQYPLADAHRLRLILAVSNKLRPLALDKLPNLRQKFGTDAVSLQAPSVWQQNAEASGNS 120
1JJIB            -----FDSLPKFDQFSSAREYREAINRIYEERNRQLSQH 50
                  : * . * * : : * . :      : : : . . . :

LIPASA-2-17      ENAVSWQDKTIANADGGDMTVRCYQQSKQNNNEGKSTDEVAMLFFHGGGFCIGDIDTHHEF 180
1JJIB            ERVERVEDRTIKGRNG-DIRVRVYQQK-----PDSPVLVYYHGGGFVICSIESHDAL 101
                  * . . : * : * * . : * * : * * * * . . * . : : : * * * * * * * * :

LIPASA-2-17      CHTVCAQTGWAVVSVDYRMAPEYPAPTALKDCLSAYAWLTEHSQSLGALPSRIVLSGDA 240
1JJIB            CRRIARLSNSTVVSVDYRLAPEHKFPAAVYDCYDATKWWAENAELRIDPSKIFVGGDA 161
                  * : . . . : * * * * * : * * : * * . * * : : : : * * * : * . : * * * *

LIPASA-2-17      GGCLAALVAQQVIKPIDALWQDNQATETDKKANDTFKKSADLPRPLAQLPLYPVTDYE 300
1JJIB            GGNLAAAVS IMARDSGEDFIKH-----QILIYPVNVNFV 194
                  ** * * * * : . . . : : : . :      * : * * * . : :

LIPASA-2-17      AEYPSWELYGEGLLLDHNDANVFNTAYTQHSGLMQSHPLISVMHGDNAHLCPYVVVAEL 360
1JJIB            APTPSLLEFGEGLWILDQKIMSWFSEQYFSREEDKFNPLASVIFADLENLPPALIITAEY 254
                  * ** : * * * : . . . : : : : : : * * * * . . * * : * : : * *

LIPASA-2-17      DILRDEGLAYAEELLQKEGVKVEYTYTVLGAPHGFINLMSVHQGLGNQNTNYI IKSFACFVQN 420
1JJIB            DPLRDEGEVFGQMLRRAGVEASIVRYRGLVHGFINYYPVLKAARDAINQIAALLVFD--- 311
                  * * * * * . : : : * : : * * . . . * . * * * * . * . . : * * : .

LIPASA-2-17      LLTSEGDEPNL 431
1JJIB            -----

```

**Figura 20: Alineamiento de secuencias entre la carboxilesterasa AFEST y la lipasa 2-17.** Se destacan los aminoácidos de la triada catalítica, siendo S: Serina (verde), D: Aspartato (amarillo) y H: Histidina (púrpura). En letras rojas se indica el codo catalítico donde se encuentra ubicada la serina.

Además, se analizó el sitio activo de la carboxilesterasa AFEST (PDB ID 1JJI), usada como molde. Se observó que éste está formado por el residuo nucleofílico **Serina 160**, la **Histidina 285** como el portador de protones y el **Aspartato 255** como residuo catalítico. Una vez conocidos estos residuos se alineó la secuencia de AFEST con la secuencia de la lipasa 2-17, se buscó cada uno de los aminoácidos de la triada catalítica presentes en la carboxilesterasa y se comparó con el aminoácido respectivo en la secuencia de la lipasa 2-17. Cada aminoácido de la triada catalítica de la carboxilesterasa AFEST, estaba alineado con el mismo tipo de aminoácido en la secuencia de la lipasa 2-17, es decir, tanto a la serina, el aspartato como a la histidina de AFEST, les correspondían una serina, un aspartato y una histidina respectivamente en la lipasa 2-17 (Figura 20).

Fue posible establecer entonces, que efectivamente la serina se encuentra en una zona muy conservada formada por 6 aminoácidos (marcados en letras rojas), y que tanto al aspartato (color amarillo) como a la histidina (color púrpura) de la carboxilesterasa le correspondía el mismo aminoácido en la secuencia de la lipasa 2-17. Por lo tanto se determinó que los aminoácidos de la triada catalítica de la lipasa 2-17 podrían corresponder a la **Serina 239** como nucleófilo, presente en la secuencia conservada GDSAGG, el **Aspartato 361** y la **Histidina 391**.

- **Alineamiento de estructuras**

El alineamiento de estructuras se basa en la comparación de la forma de las proteínas, estableciendo equivalencias en base a la conformación espacial de los átomos que las componen. Las proteínas homologas tienen regiones que conservan el mismo plegamiento y regiones en las que el plegamiento es diferente. Para pares de proteínas alejadas (identidad de residuos de aproximadamente un 20%), las regiones con el mismo plegamiento pueden comprender menos de la mitad de cada molécula. Las regiones con el mismo plegamiento en general se diferencian en su estructura a medida que aumentan las secuencias de aminoácidos divergentes.

Para determinar si existe una relación espacial entre los aminoácidos de la triada catalítica de la lipasa 2-17 y los aminoácidos de la triada catalítica de la carboxilesterasa AFEST, se realizó un alineamiento estructural de ambas proteínas mediante la herramienta "*structure comparison*" del programa Chimera. Con ello, es posible establecer todos los aminoácidos que a lo largo de la proteína podrían estar ocupando la misma distribución espacial. El programa entrega la información de la estructura secundaria en la cual se encuentra cada uno de los aminoácidos o si corresponde a un *loop*, no tiene estructura secundaria atribuida. El alineamiento se realizó en base al mayor número de aminoácidos idénticos y/o con propiedades semejantes.

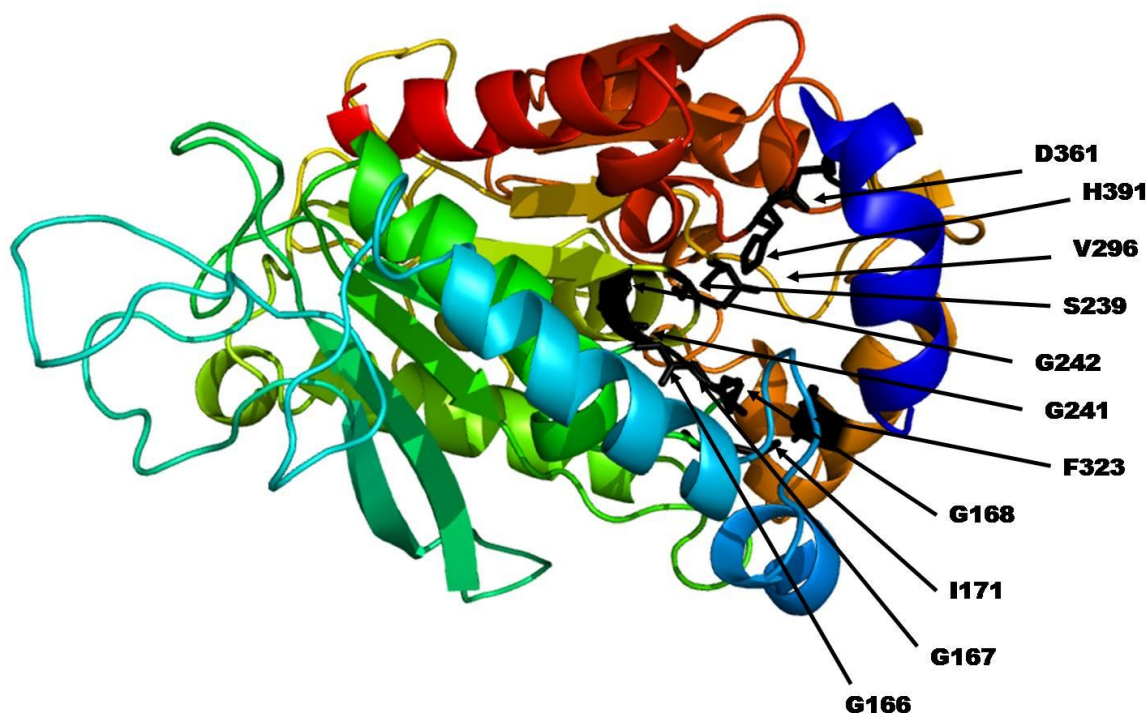


**Figura 21: Alineamiento estructural de la lipasa 2-17 y la cadena A de la lipasa AFEST.** En rosa se exhiben los aminoácidos que presentan una misma distribución espacial. Dentro del cuadrado rojo se encuentran los tres aminoácidos seleccionados como candidatos de la triada catalítica.

El alineamiento estructural (Figura 21) mostró que los aminoácidos serina 160, aspartato 255 e histidina 285 de la carboxilesterasa se encuentran alineadas con la serina 239, el aspartato 361 y la histidina 391 de la lipasa 2-17.

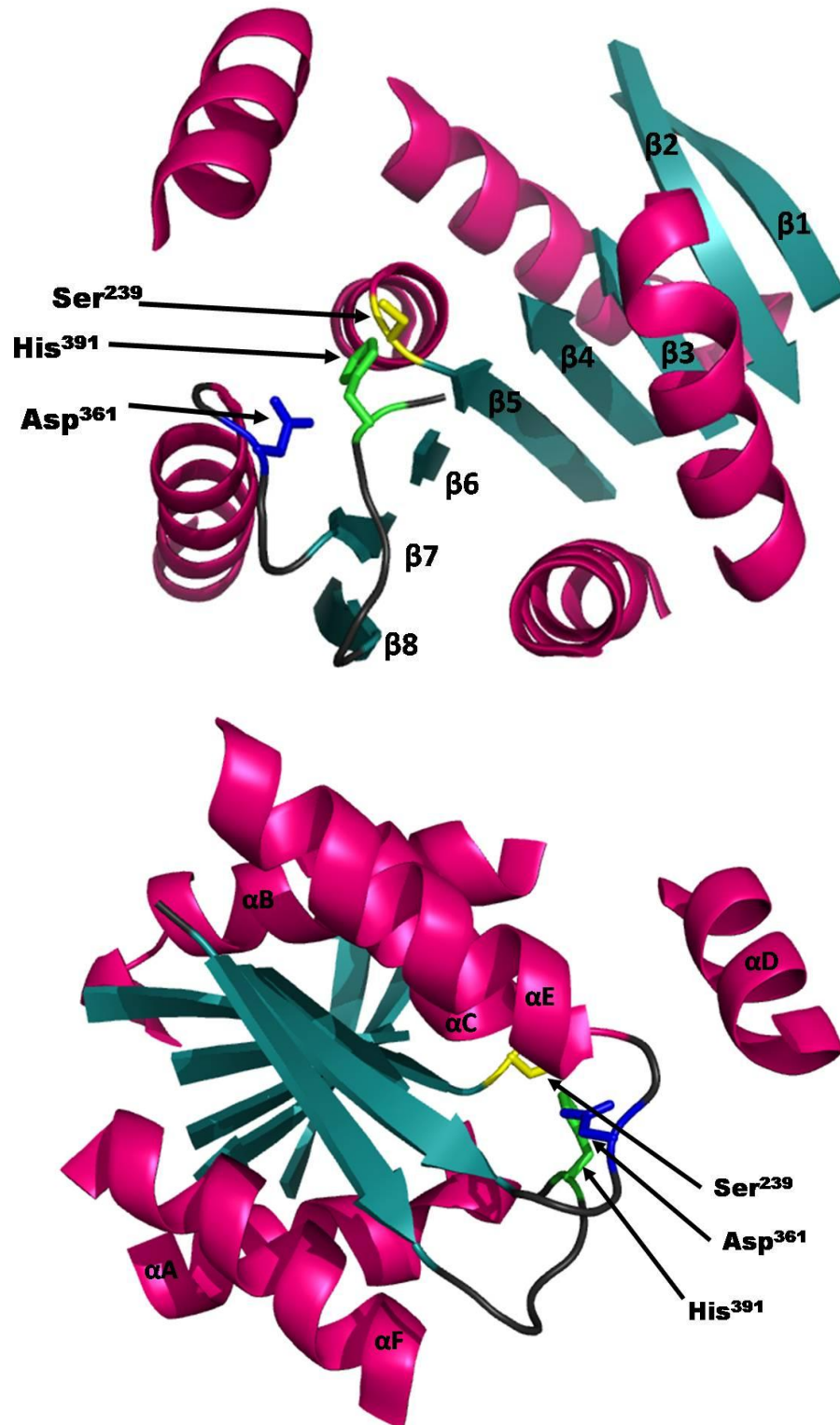
- **Análisis de las estructuras**

Mediante los programas Pymol y SPDBViewer se seleccionaron los aminoácidos que podrían estar interactuando con la triada catalítica propuesta para la lipasa 2-17 y por lo tanto que podrían estar participando en el sitio activo. Esta selección se hizo determinando todos los aminoácidos que forman enlaces con los aminoácidos de la triada propuesta y que se encontraban en un radio de 4Å. Los aminoácidos seleccionados correspondieron a G166, G167, G168, I171, S239, G241, G242, V296, F323, D361 y H391 (Figura 22).

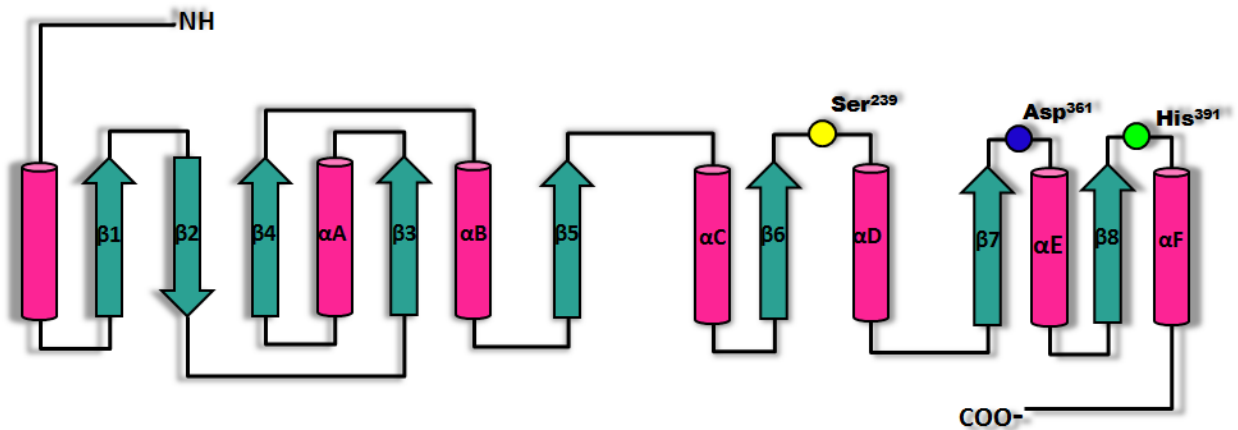


**Figura 22: Modelo tridimensional de lipasa 2-17 con aminoácidos propuestos para el sitio activo.** Figura coloreada según el espectro de colores. Extremo amino de color azul hasta extremo carboxilo de color rojo. Los aminoácidos del sitio de unión al sustrato están coloreados con púrpura. Figura realizada con el programa Pymol.

Al observar la estructura de la lipasa 2-17 es posible ver que los tres residuos candidatos se encuentran ubicados en las posiciones canónicas del plegamiento  $\alpha/\beta$  hidrolasa (Figura 23). La serina 239 se encuentra en el llamado “codo nucleofílico” entre las hojas  $\beta_5$  y  $\alpha_C$ . La cadena lateral de la histidina 391 ubicada en un *loop* de 6 residuos, que conectan la hoja  $\beta_8$  con la hélice  $\alpha_F$ , apunta con el átomo  $N^{\epsilon 2}$  hacia el  $O^{\gamma}$  de la serina 239 y se encuentra separado por una distancia de 3.10 Å. El átomo  $His391N^{\delta 1}$  está a una distancia de 3.32 Å del átomo  $O^{\delta 1}$  y de 2.26 Å del átomo  $O^{\delta 2}$  del aspartato 361. Este aminoácido se encuentra ubicado en un *loop* formado por 4 residuos (Figura 24), ubicado entre la hebra  $\beta_7$  y la hélice  $\alpha_E$ . La posición de este aspartato es posteriormente estabilizada por puentes de hidrogeno. Al comparar las distancias entre los aminoácidos de la triada catalítica de la lipasa 2-17 con las distancias de la triada de la carboxilesterasa, es posible ver que son similares. Para el caso de la carboxilesterasa, el  $N^{\delta 1}$  de la histidina se encuentra a una distancia de 2.51 Å del  $O^{\gamma}$  de la serina nucleofílica, mientras que el  $N^{\epsilon 2}$  se encuentra separado de  $O^{\delta 1}$  y de  $O^{\delta 2}$  del aspartato, por una distancia de 3.38 y 2.54 Å respectivamente. Los ángulos de torsión para la serina nucleofílica de la lipasa 2-17 correspondieron a  $\phi=72,63$  y  $\psi=-123,30$ , para el aspartato a  $\phi=-62,16$  y  $\psi=142,01$  y para la histidina a  $\phi=-158,51$  y  $\psi=136,10$ .

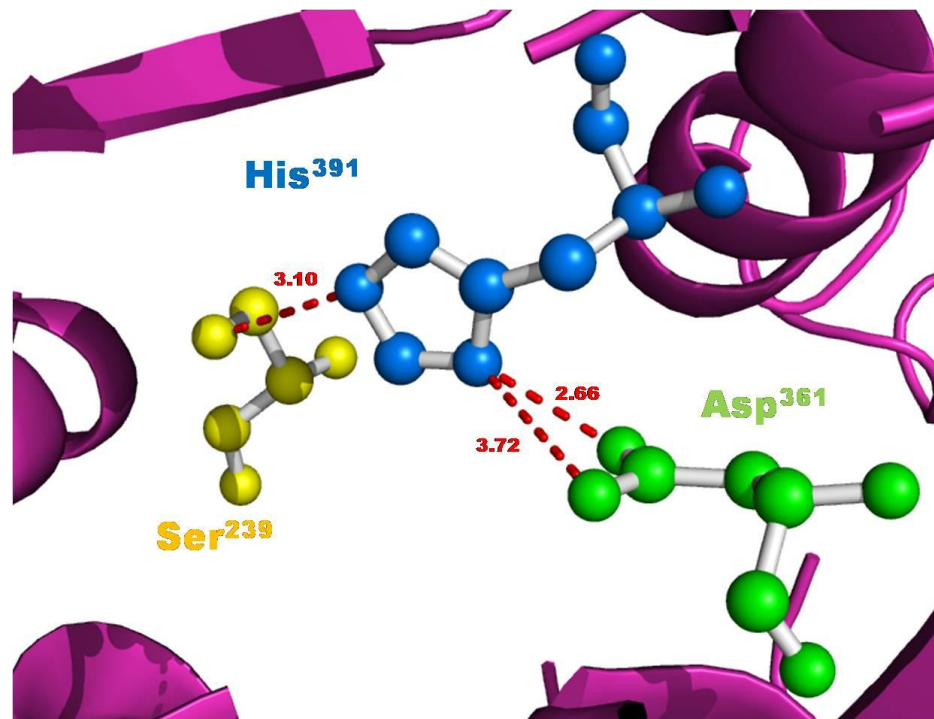


**Figura 23: Plegamiento  $\alpha/\beta$  hidrolasa.** Arriba: Modelo de la estructura central de la lipasa 2-17 visto a lo largo de la hoja  $\beta$  central. Abajo: Modelo de la estructura central de la lipasa 2-17 visto en base a las  $\alpha$  hélices. En ambos casos las hojas  $\beta$  centrales se encuentran en color celeste y las estructuras  $\alpha$  hélices en color fucsia. En amarillo se destaca la Serina, en azul el Aspartato y en verde la Histidina de la triada catalítica.



**Figura 24: Esquema general del plegamiento  $\alpha/\beta$  hidrolasa observado en la lipasa 2-17.** Las hojas  $\beta$  centrales se encuentran en color verde y las estructuras  $\alpha$  hélices en color fucsia. En amarillo se destaca la Serina, en azul el Aspartato y en verde la Histidina de la triada catalítica.

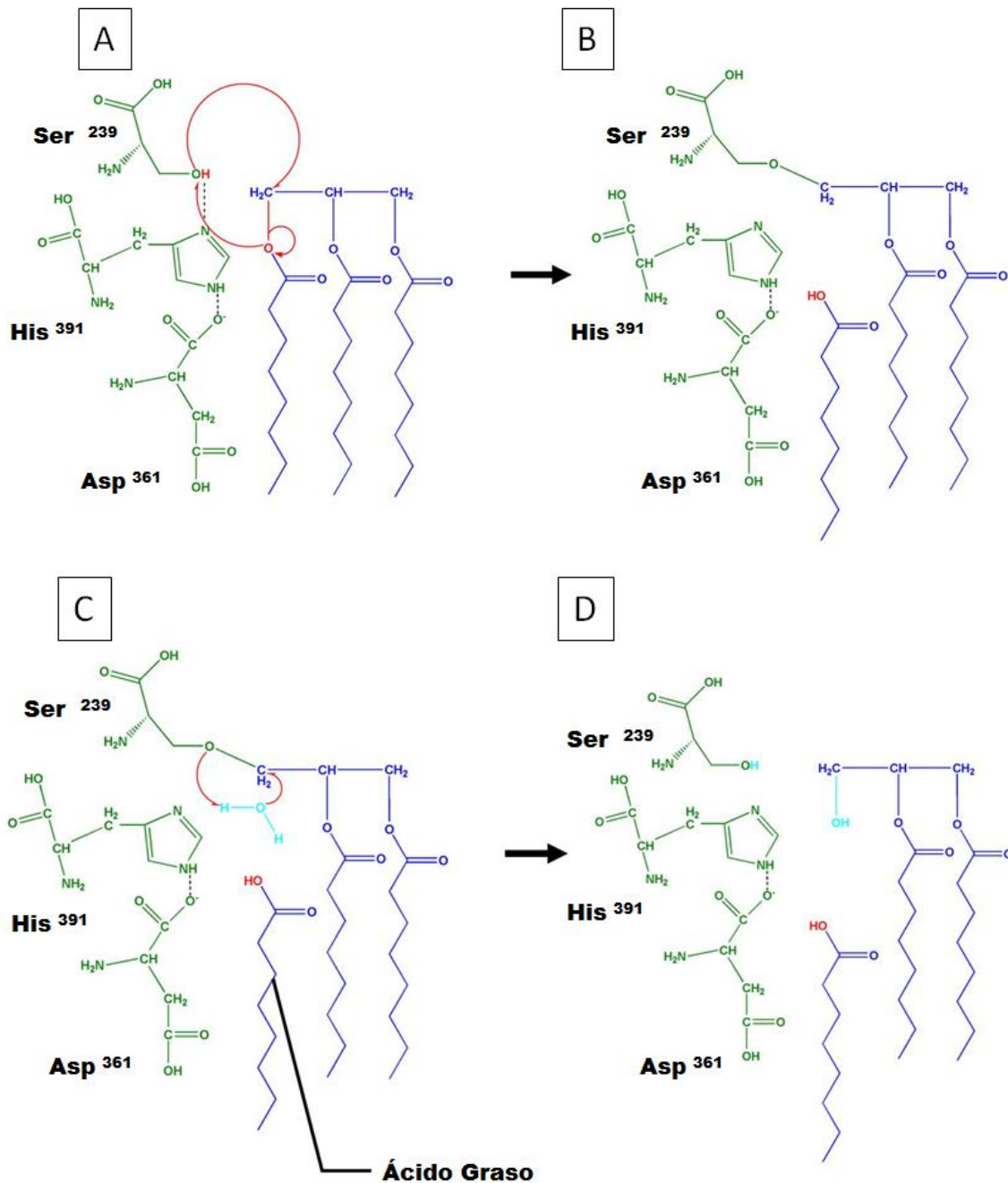
Se sabe que el ciclo catalítico de las lipasas se inicia con la activación de la serina por la histidina que se encuentra ubicada junto a ella y el ataque nucleofílico al grupo carbonil del sustrato realizado por el átomo de oxígeno del grupo hidroxilo del residuo serina. Así, se forma un intermediario tetraédrico transitorio, que se estabiliza por residuos de la cavidad oxianiónica. Los pasos de este mecanismo se encuentran ilustrados en la Figura 26.



**Figura 25: Triada catalítica propuesta para la lipasa 2-17.** En púrpura se exhibe el modelo tridimensional de la lipasa 2-17. Los aminoácidos señalados corresponden a la triada catalítica, Ser 239 (en amarillo), Asp 361 (en verde) e His 391 (en morado). Figura realizada con el programa Pymol.



Como es posible ver en la Figura 25, es fundamental que los tres aminoácidos de la triada catalítica se encuentren espacialmente cercanos para poder llevar a cabo la reacción lipolítica. Por ello, para comprobar que los aminoácidos **Ser 239**, **His 391** y **Asp 361** pertenecían efectivamente a la triada catalítica de la lipasa 2-17, se visualizaron en la estructura tridimensional establecida para la lipasa 2-17.



**Figura 26: Mecanismo catalítico de lipasas.** En el esquema se indican los pasos involucrados en la catálisis enzimática de las lipasas: **A:** unión del sustrato. **B:** formación del intermediario tetraédrico. **C:** formación de intermediario covalente. **D:** liberación del producto final y regeneración del catalizador (basado en Jaeger y col., 1999).

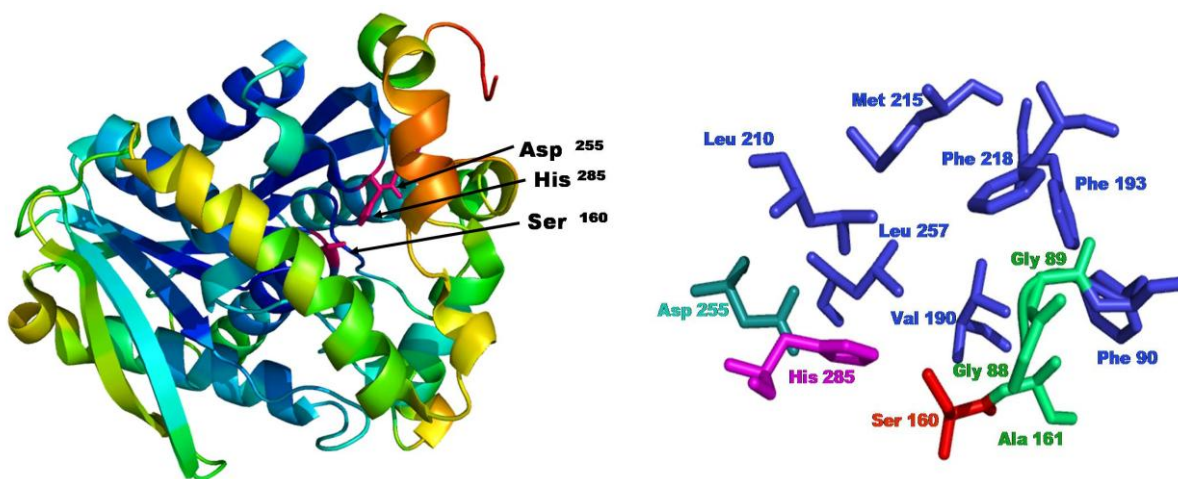
Debido a la proximidad espacial de los tres aminoácidos propuestos, fueron establecidos como posibles candidatos para la triada catalítica de la lipasa 2-17 y se procedió a estudiar a los aminoácidos que podrían estar interactuando con la serina 239, el aspartato 361 y la histidina 391.

### 3.4.1.3 Comparación de las estructuras de lipasa 2-17 con la termófila AFEST

Para predecir correctamente la estructura de la lipasa 2-17 y conocer su relación con la función, se realizó un análisis comparativo con la estructura tridimensional de la proteína molde, la carboxilesterasa AFEST. Lo primero fue analizar la triada catalítica de la carboxilesterasa, comprar los aminoácidos del sitio activo de cada una de las proteínas y finalmente realizar una comparación directa de los aminoácidos de la triada catalítica.

- **Descripción de la triada catalítica de la carboxilesterasa 1JJJ**

El sitio activo de la lipasa AFEST consiste en una tríada catalítica clásica que cuenta con Ser 160 como nucleófilo, His 285 como portador de protones y Asp 255 que participa en la formación de la red de cargas. Todos estos residuos están ubicados en forma canónica en el clásico plegamiento  $\alpha/\beta$  hidrolasa. La serina 160 es parte de la secuencia conservada GDSAGG y se encuentra ubicada en el codo nucleofílico entre la hoja  $\beta$ 5 y la  $\alpha$  hélice  $\alpha$ 7 (Figura 27).



**Figura 27: Modelo y estereodiagrama del sitio catalítico de AFEST.** (Modificado de De Simone y col., 2001). **Izquierda:** Modelo tridimensional de la lipasa AFEST. En púrpura los aminoácidos de la triada catalítica. **Derecha:** Estereodiagrama del sitio catalítico de la lipasa AFEST. En celeste, púrpura y rojo los aminoácidos de la triada catalítica Asp 255, His 285 y Serina 160 respectivamente. En azul, el resto de los principales aminoácidos del sitio activo.

La serina nucleofílica se encuentra ubicada en la parte inferior de un túnel hidrofóbico. La entrada de este túnel está rodeada de dos  $\alpha$  hélices ( $\alpha 8$  y  $\alpha 9$ ), la hélice G3 y los *loops* formados por los aminoácidos Val194-Thr197, Phe203-Asp211 y Phe223-Arg225. Las regiones de los *loops* formados por los aminoácidos Val189-Asn191 y Gly87-Phe90 actúan como fronteras internas de este túnel.

- **Comparación del sitio catalítico de la lipasa 2-17 con el de la carboxilesterasa 1JJJ**

Finalmente se analizaron y compararon todos los aminoácidos correspondientes al sitio catalítico de la carboxilesterasa 1JJJ. Para ello se utilizó el programa Pymol y se seleccionaron los aminoácidos Ser 160- Asp 255- His 285, que corresponden a la triada catalítica y los aminoácidos Gly 88- Gly 89- Phe 90- Ala 161- Phe 193- Leu 210- Met 215- Phe 218- Leu 257, que corresponden a los aminoácidos que acompañan a la triada en el sitio activo y participan directamente en la catálisis del sustrato (De Simone y col., 2001) (Figura 28).

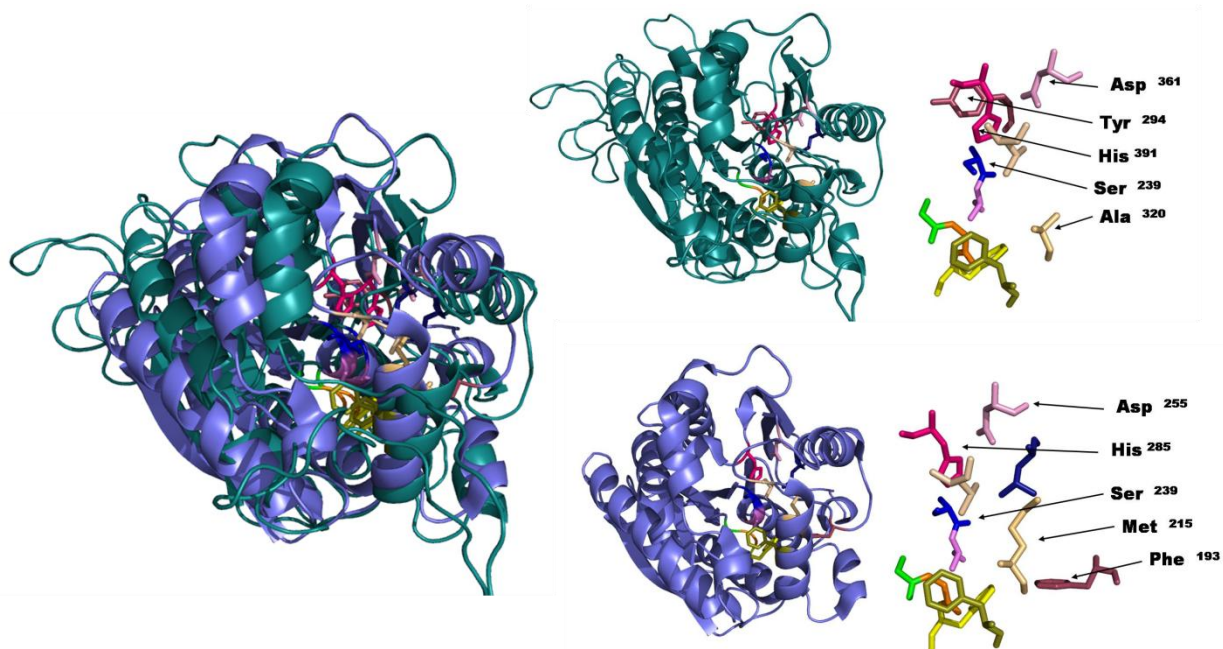
Utilizando la información obtenida previamente y el programa Pymol se realizó una selección de todos los aminoácidos de la lipasa 2-17 que acompañan a la triada catalítica, con el fin de analizar las propiedades de cada uno de estos aminoácidos y compararlos con los descritos para la carboxilesterasa 1JJJ. Se determinó que los aminoácidos del sitio activo de la lipasa 2-17 que acompañan a la triada catalítica formada por Ser 239 - Asp 361- His 391, corresponden a Gly 167 - Gly 168 - Phe 169 - Ala 237 - Tyr 294 - Leu 315 - Ala 320 - Phe 323- Leu 363. Es posible observar que los únicos aminoácidos que no son correspondientes en cada uno de los modelos son Phe 193 - Met 215 presentes en la carboxilesterasa y los aminoácidos Tyr 294 - Ala 320, presentes en la lipasa 2-17 (Figura 28).

Si bien, tanto la fenilalanina presente en la carboxilesterasa como la tirosina de la lipasa 2-17 son aminoácidos neutros, la fenilalanina es un aminoácido apolar mientras que la tirosina es casi idéntica, pero presenta un grupo polar sin carga. Sus pesos moleculares son relativamente similares (Phe 165.19 y Tyr 181.19) y su presencia en una proteína oscila en un 3.91% para la fenilalanina y un 3.25% para la tirosina, pero la tirosina de la lipasa 2-17 se encuentra ubicada cerca de la histidina catalítica, mientras que la fenilalanina se encuentra cercana a la fenilalanina 193 (correspondiente a la 323 de la lipasa 2-17), que provoca un angostamiento en la cavidad del sitio catalítico en la carboxilesterasa.

Por otro lado, también es posible observar que, en la secuencia de la carboxilesterasa, se encuentra participando en el sitio activo una metionina, mientras que en la lipasa lo hace una alanina. Ambos aminoácidos son neutros y apolares, pero a pesar de que el grupo metilo de la alanina es muy poco reactivo, a veces puede desempeñar una función en el reconocimiento del sustrato o especificidad, interactuando con otros grupos apolares.

Mediante una comparación de los aminoácidos del sitio catalítico utilizando el programa SPDBViewer, se estableció que los puentes de hidrógeno formados en la carboxilesterasa correspondían

a His285 – Ser160 (2.51 Å) y Asp255 – Leu257 (2.78 Å), mientras que los puentes formados en la lipasa 2-17 fueron 3: Ser 239-His391 (3.1 Å), Asp 361- Leu 363 (3.01 Å) y Ala 320- Phe 323 (3.3 Å). Este último puente se forma por la presencia de la Alanina que se encuentra en la lipasa 2-17 pero que no está presente en la carboxilesterasa (cuyo aminoácido respectivo estructuralmente es la metionina comparada previamente). Este enlace acerca a la fenilalanina 323 de la lipasa 2-17 hacia la alanina 320 (fenilalanina 193 respectiva en la carboxilesterasa) y podría estar condicionando los grados de libertad del sitio activo y participando en la selectividad de la enzima por su sustrato.



**Figura 28: Comparación de los aminoácidos del sitio activo de la carboxilesterasa 1JJI y de la lipasa 2-17. Izquierda:** Superposición de la carboxilesterasa 1JJI y la lipasa 2-17 con cada uno de los aminoácidos del sitio activo marcados con un color distintivo. **Arriba:** Carboxilesterasa con los aminoácidos del sitio activo Ser 239 - Asp 361- His 391 - Gly 167 - Gly 168 - Phe 169 - Ala 237 - Tyr 294 - Leu 315 - Ala 320 - Phe 323- Leu 363 destacados en colores. **Abajo:** Lipasa 2-17 con los aminoácidos del sitio activo Ser 239 - Asp 361- His 391 - Gly 167 - Gly 168 - Phe 169 - Ala 237 - Tyr 294 - Leu 315 - Ala 320 - Phe 323- Leu 363 destacados con los mismos colores que la carboxilesterasa.

Con el programa SPDBViewer y herramienta “*Neighbors of selected residues*”, se establecieron todos los aminoácidos cercanos al sitio catalítico con el objetivo de ampliar el espectro de aminoácidos para conocer las propiedades del sitio activo de la lipasa 2-17. Los aminoácidos cercanos a la triada catalítica de la carboxilesterasa 1JJI con 6Å de distancia desde la triada fueron 39 (Tyr10 – Gly87 – Gly88 – Gly89 – Phe90 – Gly158 – Asp159 – Ser160 – Ala161 – hGly162 – Gly163 – Asn164 – Leu165 – Ile187 – Tyr 188 – Pro189 – Val190 – Ser199 – Leu200 – Phe203 – Leu207 – Trp208 – Ile209 – Leu210 – Thr251-

Ala252 – Glu253 – Tyr254 – Asp255 – Pro256 – Leu257 – Arg258 – Asp259 – Val283 – Leu284 – His285 – Gly286 – Phe287 – Tyr290), mientras que los aminoácidos cercanos a la triada catalítica de la lipasa 2-17 con los mismos 6Å de distancia desde la triada, correspondieron a 26 (Asp238 - Ser239 - Ala240 - Leu293 – Tyr294 – Pro295 – Ser305 – Trp306 – Tyr309 – Leu313 – Leu314 – Leu315 – Ala358 – Glu359 – Leu360 – Asp361 – Ile362 – Leu363 – Arg364 – Asp365 – Ala389 – Pro390 – His391 – Gly392 – Phe393 – Ile394). Esto mostró que el sitio catalítico de la carboxilesterasa comprende un universo de aminoácidos mayor que el observado en la lipasa 2-17, lo que podría verse reflejado en los tipos de sustrato que ambas enzimas aceptan.

Finalmente, con el objetivo de corroborar los análisis anteriores, se realizó una superposición estructural de la lipasa 2-17 y la carboxilesterasa AFEST, esta vez dejando fijos los aminoácidos de la triada catalítica. Esto permite observar el ordenamiento espacial de todo el resto de la proteína y por lo tanto de los otros aminoácidos del sitio activo y evaluar la variación con respecto a la lipasa molde. Los resultados revelaron que si bien ambas proteínas actúan en rangos de temperaturas muy diferentes, sus estructuras y sus centros activos son muy similares en términos de la composición aminoacídica, pero el ordenamiento espacial que tienen, difiere fundamentalmente en el número de aminoácidos que componen el sitio activo.

#### **3.4.1.4 Comparación de las estructuras de lipasa 2-17 con la psicrófila 2ORY**

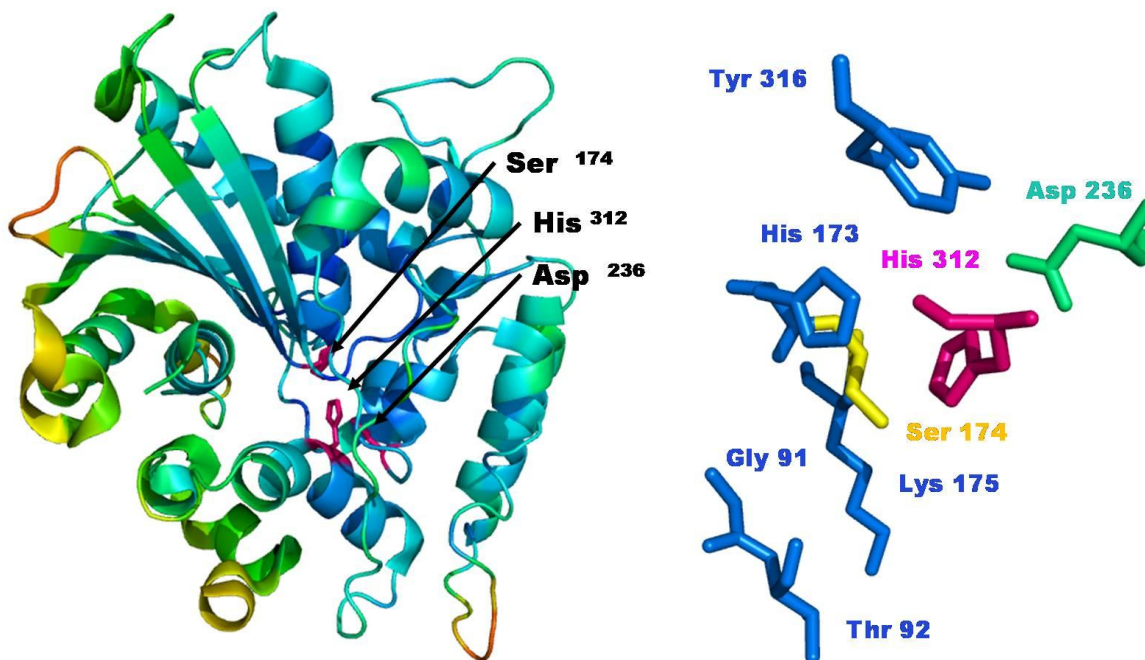
Para conocer más a fondo la relación entre la estructura y la función se realizó una comparación de la lipasa 2-17 con otra lipasa psicrófila que también proviene de una bacteria. Al igual que lo realizado con la carboxilesterasa (sección 3.4.1.3), se analizó la triada catalítica de la lipasa M37 (2ORY), se compararon los aminoácidos del sitio activo de ambas proteínas y se estudiaron los aminoácidos de ambos sitios catalíticos.

- **Descripción de la triada catalítica de la psicrófila 2ORY**

La tríada catalítica de la lipasa M37 está formada por los residuos Ser174, Asp236, e His312 y se encuentra ubicada en una región altamente conservada. Su sitio activo está cubierto por una  $\alpha$  hélice ( $\alpha 3$ ) que actúa como tapa, similar a las estructuras observadas en otras lipasas bacterianas.

El sitio activo de la lipasa M37 (Figura 29) se encuentra completamente inaccesible para el sustrato cuando la conformación de la enzima es cerrada. Bajo la tapa, se puede observar una cavidad

que se encuentra rodeada por los residuos de la cadena  $\beta 2$  (Phe68, Asn69 y Ala71),  $\alpha 4$  (Ser128, Thr129 y Thr135),  $\alpha 3$  (Trp100)  $\alpha 6$  (Lys175, Ala178 y Leu179) y varios *loops* (Leu101, Asp104, Pro209, y Trp242).



**Figura 29: Modelo y diagrama del sitio catalítico de la lipasa M37 (2ORY).** (Modificado de Jung S.K. y col., 2008). **Izquierda:** Modelo tridimensional de la lipasa M37. En púrpura los aminoácidos de la triada catalítica. **Derecha:** Diagrama del sitio catalítico de la lipasa M37. En celeste, púrpura y rojo los aminoácidos de la triada catalítica Asp 236, His 312 y Serina 174 respectivamente. En azul, el resto de los principales aminoácidos del sitio activo.

Bajo esta cavidad, el sitio activo de la lipasa M37 está compuesto principalmente por los aminoácidos Gly91, Thr92, His173, Ser174, Lys175, Asp236, His312 y Lys316. Se estableció que los puentes de hidrógeno formados por estos aminoácidos eran 4 y correspondían a: His312 – Ser174 (2.93 Å), Asp104 – Ser128 (2.64 Å), Asp104 – Leu101 (3.20 Å), Lys175 – Ala178 (2.90 / 3.12 Å).

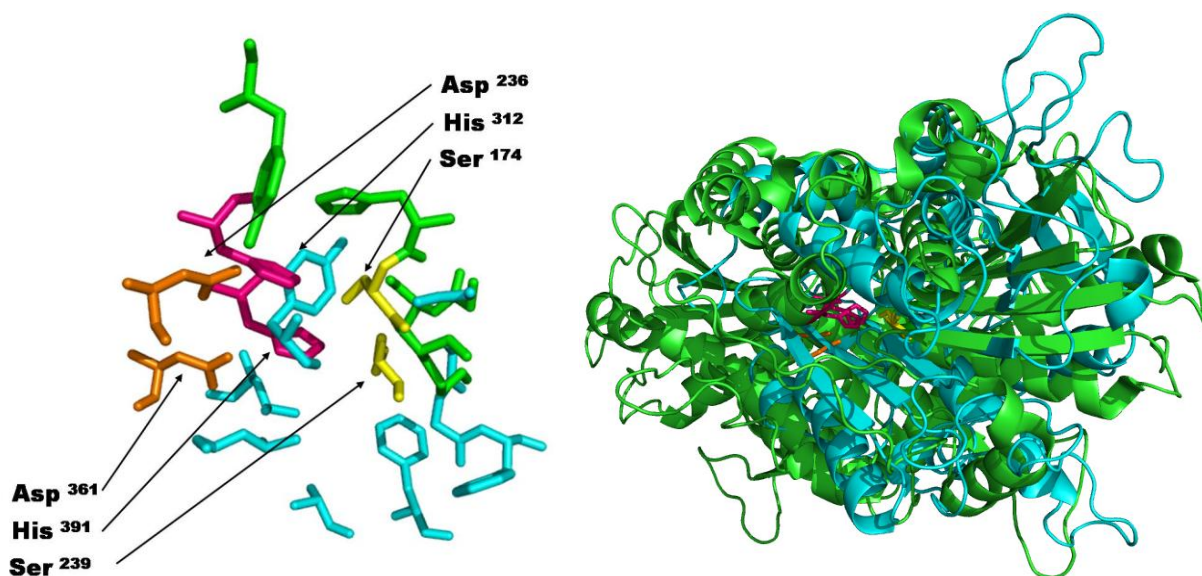
- **Comparación del sitio catalítico de la lipasa 2-17 y la lipasa 2ORY**

Mediante el programa Pymol se seleccionaron los aminoácidos determinados por Jung y col., como los aminoácidos del sitio activo de la lipasa M37 que acompañan catalíticamente a la triada Ser174 – Asp236 – His312. Estos son: Gly91 – Thr92 – His173 – Lys175 y Tyr316 (Figura 30).

Estos aminoácidos fueron comparados con los aminoácidos seleccionados para la lipasa 2-17 y fue posible ver que no existe una correlación entre ambos sitios catalíticos, a pesar de las similitudes en sus propiedades y en las especificidades por ciertos sustratos.

Al ampliar el campo de aminoácidos cercanos al sitio catalítico a 6 Å fue posible observar sin embargo, que ambas enzimas presentan un universo similar de aminoácidos tanto en cantidad, como en

tipo y composición. Para el caso de la lipasa 2-17, los aminoácidos vecinos a la triada catalítica correspondieron a 23 (descritos en la sección anterior), mientras que para la lipasa M37 fueron 25: Tyr14 – Tyr17 – Thr92 – Asn93 – Pro94 – Val95 – Ile94 – Trp100 – Gly172 – His173 – Ser174 – Lys175 – Gly176 – Gly177 – Ala178 – Leu179 – Phe206 – Ala207 – Gly208 – Pro209 – Ala232 – Asn233 – Ser234 – Leu235 – Asp236 – Ile237 – Val238 – Pro239 – Tyr240 – Ley250 – Ile253 – Tyr254 – Gln307 – Ala308 – Ala309 – Tyr310 – Gln311 – His312 – Val313 – Val314 – Gly315 – Tyr316.



**Figura 30: Comparación de los aminoácidos sitio activo lipasa 2ORY y la lipasa 2-17.** Izquierda: Comparación de los aminoácidos del sitio activo de ambas enzimas, en base a la superposición de la triada catalítica. En celeste aminoácidos de la lipasa 2-17, en verde aminoácidos de la lipasa M37. Triada catalítica coloreada en amarillo, fucsia y anaranjado para los aminoácidos serina, histidina y aspartato, respectivamente. Derecha: Superposición de la lipasa M37 y la lipasa 2-17 con cada uno de los aminoácidos de la triada catalítica marcados con el color descrito previamente.

Por otro lado, ambas enzimas coinciden en la presencia de 11 aminoácidos: Tyr, Pro, Ile, Trp, His, Ser, Phe, Ala, Gly, Leu y Asp. Si bien existen diferencias para el resto de los aminoácidos, la mayoría de los aminoácidos comparados son aminoácidos con propiedades similares (polares no cargados o con carga positiva), pero que podrían estar participando en las diferencias en la selectividad de sustrato de cada una de las enzimas.

En ninguno de los sitios catalíticos de las enzimas comparadas se observó la presencia de cisteínas que en caso de estar presente, podrían formar enlaces disulfuros, dificultando o impidiendo la entrada del sustrato al sitio activo.

### 3.4.2 Estudio de la interacción lipasa nativa-triacilglicerol mediante *Docking* Molecular

El acoplamiento molecular o *docking* es una técnica de mecánica molecular que se utiliza para predecir energías y modos de enlaces entre ligandos y proteínas, por lo tanto es posible estimar la estructura de los complejos moleculares formados entre dos o más moléculas.

Según la caracterización realizada para la lipasa 2-17, el sustrato preferente son triacilgliceroles de 6 carbonos, por lo tanto para realizar el *docking* molecular de la lipasa 2-17 (molecula receptora) se seleccionaron 4 tipos de triacilgliceroles (ligandos) formados por los ácidos grasos Caproico (de 6 carbonos), Caprílico (de 8 carbonos), Cáprico (de 10 carbonos) y Palmítico (de 16 carbonos), con el objetivo de evaluar el acoplamiento de cada uno de ellos (Figura 31).

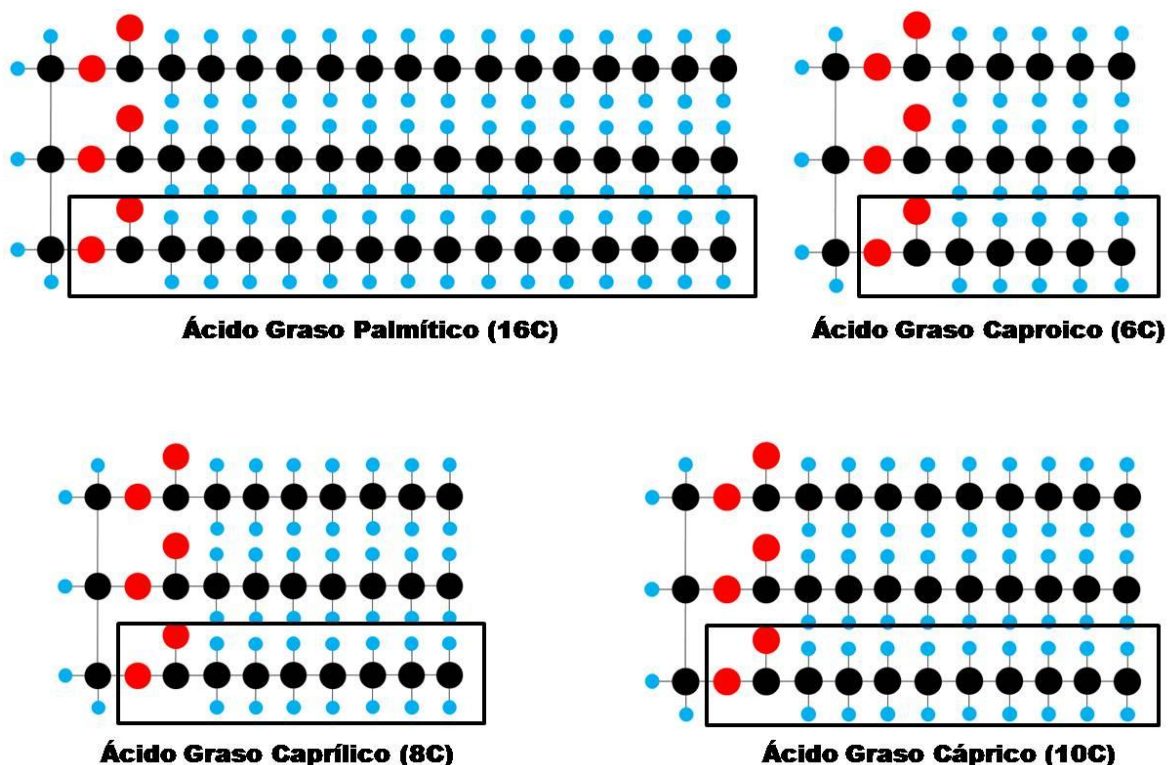


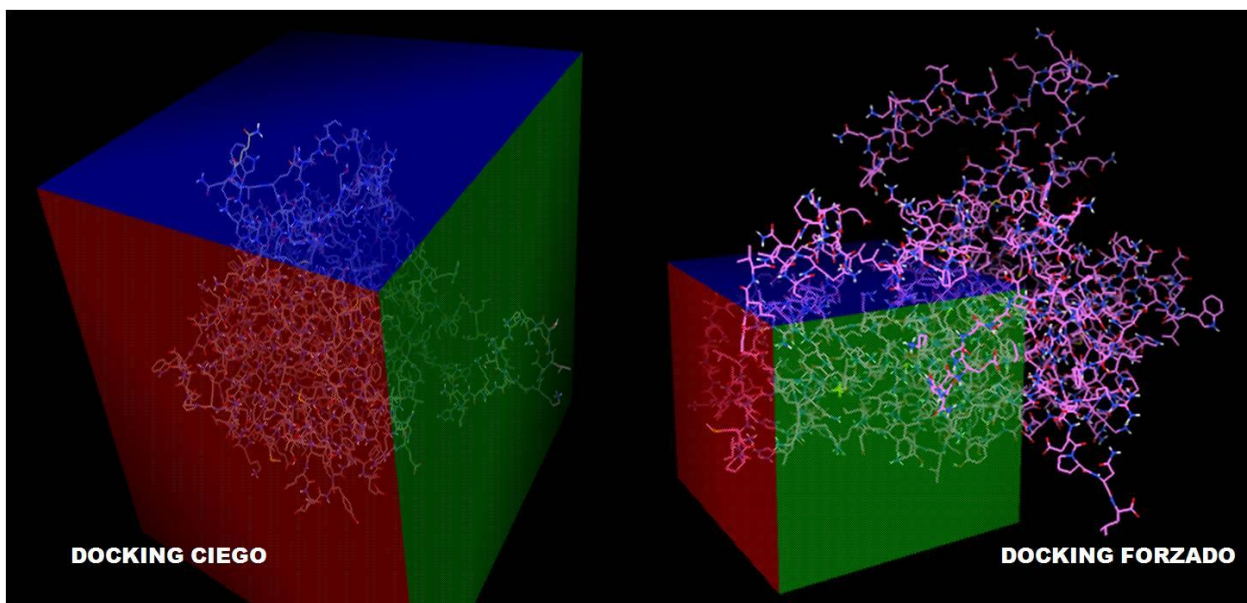
Figura 31: Esquema representativo de los triacilgliceroles usados para el *docking* molecular. Triacilgliceroles con ácidos grasos palmítico (16C), cáprico (10C), caprílico (8C) y caproico (6C), marcados con rectángulo negro.

Para comenzar el proceso de *docking*, lo primero fue preparar los archivos tanto de la proteína como de los ligandos. Para el caso de los triacilgliceroles, en primera instancia, éstos se definieron como moléculas rígidas, debido a lo cual fue importante chequear que el número de rotámetros posibles se encuentre en cero. En el caso del modelo de la lipasa 2-17, se le agregaron los hidrógenos polares y posteriormente las cargas parciales y se guardó el archivo en formato *pdbqt*.



El modelo tridimensional de la lipasa 2-17 obtenido previamente, se utilizó para llevar a cabo el análisis de *docking* molecular y de esta forma estudiar la unión a los cuatro tipos de triacilgliceroles nombrados previamente. El modelamiento de la lipasa 2-17 entregó un modelo cuya conformación fue con el sitio activo descubierto (conformación abierta) y con las posibles estructuras tipo tapa (descritas en el capítulo 5) posicionadas fuera de la entrada al sitio activo, lo que permitiría dejarlo accesible al sustrato. Esta conformación abierta de la lipasa 2-17 se obtuvo debido a que el molde utilizado para obtener su estructura tridimensional, presentaba justamente una conformación abierta. Este fue uno de los criterios utilizados a la hora de escoger un modelo (además de los parámetros entregados por las herramientas de Modeller y FUGUE) ya que al estar en conformación abierta, es posible realizar estudios de interacción con sustratos, inhibidores y en algunos casos cofactores.

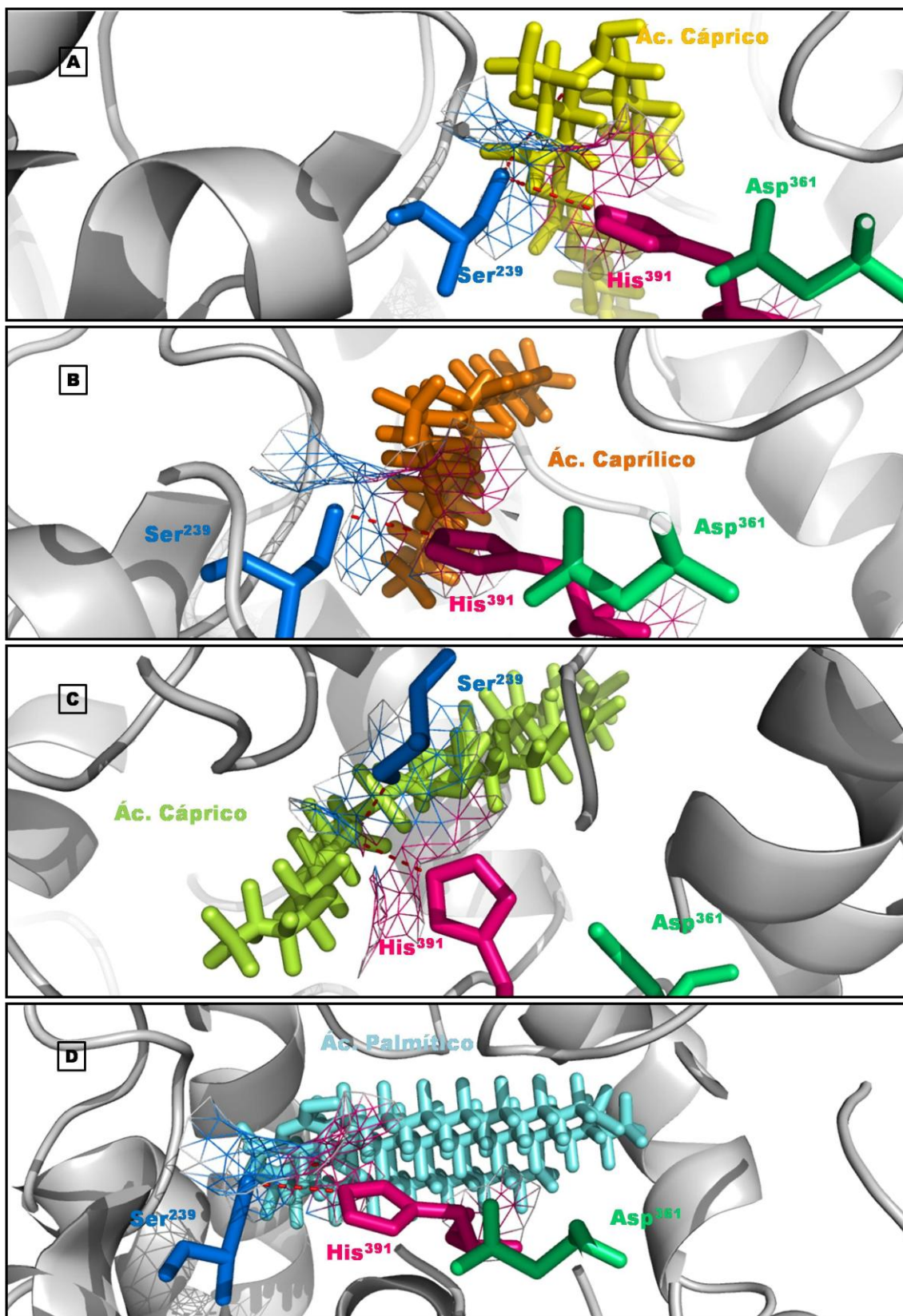
Con el fin de estudiar la unión de la lipasa a los diferentes sustratos, se utilizaron dos programas para llevar a cabo el análisis: Autodock y Vina. Se realizaron dos tipos de *docking* molecular: el *docking* ciego, que consiste en definir un espacio de búsqueda suficientemente grande de modo de permitir la unión del ligando en cualquier lugar de la lipasa, el *docking* forzado, que define un espacio de búsqueda limitado. En este caso, se limitó el espacio de búsqueda al área de interacción observada en otros dominios de unión de las lipasas EC 1.3.3.1 (Ji y col., 2010; Juhl y col., 2009), y en los aminoácidos seleccionados como posibles candidatos para la triada catalítica.



**Figura 32: Representación *Docking* Molecular** Izquierda: *Docking* Ciego. Los ejes y el tamaño del cubo son definidos en base a la molécula completa. Derecha: *Docking* Forzado: La posición del cubo y su tamaño son definidos en base a la zona candidata como sitio activo una vez realizado el primer acercamiento del *docking* ciego en la molécula completa.

La elección del tamaño del espacio de búsqueda se basó en que éste fuese suficientemente grande con el objetivo de contener la totalidad de átomos de la lipasa en estudio (receptor). En base al primer criterio se definió que la caja de búsqueda fuese de al menos 6Å x 35Å x 25Å (caja de 6 x 35 x 25, con puntos espaciados cada 1Å). Para el caso del *docking* ciego, la caja de búsqueda se centró en la lipasa, mientras que en el caso del *docking* forzado se centró de manera tal de limitar el espacio de búsqueda a la zona de interés, particularmente los aminoácidos Ser239, Asp361 e His391. Se llevó a cabo 4 simulaciones, correspondientes a la combinación de la lipasa 2-17 con 4 tipos de sustratos, en particular triacilgliceroles de 6, 8, 10 y 16 carbonos en la cadena lateral del ácido graso. Cada una de las simulaciones se realizó tanto en Autodock como en Vina. Ambos programas utilizan el mismo tipo de archivos de entrada para describir al receptor y al ligando, lo que facilita el uso de ambos programas y la comparación de sus resultados. Las funciones de puntuación tienen parámetros similares, pero ponderados de manera distinta. Una diferencia clave entre los programas es el algoritmo de búsqueda local. Autodock utiliza búsqueda estocástica que genera conformaciones al azar, Vina en cambio, calcula un gradiente mientras busca un óptimo local (Chang y col., 2010). Cuando se trabaja con *docking* rígido, el objetivo es mantener al ligando completamente rígido durante la etapa de orientación para explorar los algoritmos de correspondencia y de minimización. En el estudio del acoplamiento con *docking* flexible, se permite que el ligando pueda reorganizarse estructuralmente en respuesta a la molécula receptora, en este caso la lipasa. Así, se evaluó la orientación preferente de la lipasa sobre el sustrato, mediante un análisis de *docking* molecular, llevado a cabo con los programas Autodock y Vina. La preparación de los archivos de entrada fue similar para ambos casos, salvo porque en el caso de Autodock se requiere calcular los mapas de búsqueda. Para todos los casos, tanto *docking* ciego como forzado, las zonas de acople correspondieron a las definidas como sitio activo y los aminoácidos interactuantes fueron justamente los de la triada catalítica.

En el primer caso, el *docking* realizado para el sustrato preferente de la lipasa 2-17, triacilgliceroles con cadenas de 6 carbonos, el *docking* entregó energías de interacción consistente con las informadas para otras lipasas (Luo y col.,2006). La orientación y los enlaces formados entre la triada y la molécula de sustrato muestran que tal como se predice para las lipasas de este tipo, cuando la molécula de triacilglicerol alcanza el sitio activo en la parte inferior de la hendidura hidrofóbica, entra en contacto con la triada catalítica que lleva a cabo la reacción de hidrólisis. La triada está orientada de tal manera que uno de los oxígenos terminales del aspartato se une al hidrógeno de la Histidina (Figura 33). Esto mantiene en posición a la His391 para que el hidrógeno de su siguiente nitrógeno aromático forme un enlace con la Serina 239. Aquí se produce el ataque nucleofílico de la cadena del la Serina contra el primer carbono del glicerol.



**Figura 33: Interacción Triada catalítica lipasa 2-17 y triacilgliceroles.** A: Cadena de ácido graso de 6C (amarillo). B: Cadena de ácido graso de 8C (anaranjado). C: Cadena de ácido graso de 10C (verde claro). D: Cadena de ácido graso de 16C (celeste). Triada catalítica compuesta por serina 239, aspartato 361 e histidina 391 representada en azul, verde y rosa respectivamente.

El oxígeno que participó en el enlace éster se convierte en un “grupo saliente” y es protonado por el hidrógeno de la Serina 239, que fue retirado del grupo hidroxilo para neutralizar su carga negativa. En la Figura 33 es posible ver el enlace formado por la histidina 391 con la serina 239 y a su vez el enlace que forma la serina con el sustrato. Posteriormente se debe formar y eliminar el diglicérido que contiene el glicerol. Para eliminar la molécula de glicerol, un hidroxilo se une al carbono y un protón permite la reprotonación del oxígeno de la Serina. Esto es gracias a la incorporación de una molécula de agua. Esta molécula, a través de la interacción del hidrógeno, es llevada a su forma disociada. El hidroxilo tiene un gran momento dipolar y actúa como un nucleófilo en el primer carbono. El oxígeno de la serina es esta vez, el grupo saliente. La cadena lateral cargada negativamente de la serina se protona rápidamente por el ión de hidrógeno proveniente del agua. Se forma entonces el otro producto diglicérido. El enlace del hidrógeno entre el nitrógeno aromático de la histidina y el grupo hidroxilo de la serina se vuelve a formar ahora con un nuevo hidrógeno y la enzima se encuentra con su triada catalítica en su orientación original, lista para otra reacción. Para el caso del *docking* realizado con el sustrato de 8 carbonos (Figura 33), fue posible establecer que debido al aumento en el largo de las cadenas, no se realiza la correcta interacción del sustrato con la triada catalítica. En este caso, podría haber una diferencia de tamaño entre la cavidad oxianiónica, las hendiduras sn-1, sn-2 y sn-3 y las cadenas de la molécula del triacilglicerol. En un comienzo hay interacción del oxígeno del aspartato con el hidrógeno de la histidina lo que permite que el átomo de nitrógeno forme un enlace con la serina, pero debido a la lejanía con el carbono del sustrato, la serina no logra realizar el ataque nucleofílico correcto y la energía requerida para llegar a este carbono implicaría realizar movimientos estructurales en varias partes de la enzima, observándose un movimiento más complejo que el realizado por la serina en condiciones favorables.

Posteriormente se estudió la interacción de la lipasa 2-17 con un sustrato de cadena aun mayor, triacilglicerol con cadenas de 10 carbonos. El acoplamiento de estas dos moléculas mostró que el aumento del largo de la cadena forzó a un cambio conformacional más complejo en la enzima. Las cadenas aciladas no encajaron de manera correcta provocando que tanto la serina como la histidina catalíticas interactuaran con uno de los carbonos del sustrato, interacciones que sin embargo, no permiten correctamente la hidrólisis de este sustrato (Figura 33).

Finalmente se estudió las interacciones con un sustrato de cadena excesivamente larga (15 carbonos), para determinar la influencia del sustrato en la organización espacial de los átomos y las energías de interacción (Figura 33). El *docking* con el triacilglicerol de 16 carbonos reveló un reordenamiento de toda la estructura. Se observaron cambios en el posicionamiento de algunas hélices y a pesar de la formación de un enlace entre la histidina y la serina catalíticas, las diferencias en las energías de interacción indican que no habría hidrólisis de este sustrato. El *docking* ciego de una

molécula muy pequeña de 2 carbonos (datos no mostrados), no entregó mayoritariamente información sobre el acoplamiento en el sitio activo, pero al realizar el *docking* forzado éste mostró nuevamente una posible interacción de la histidina con la serina catalítica. A diferencia de los otros sustratos, la falta de interacción posiblemente se deba a que las cadenas de los ácidos grasos no son estabilizadas en las hendiduras y por lo tanto no estabilice la molécula de sustrato en el sitio activo. Otros estudios de dinámica molecular en algunas lipasas avalan esta hipótesis (Ji y col., 2010; Juhl y col., 2009).

Los resultados entregados por el *docking* molecular concuerdan los resultados experimentales obtenidos para la especificidad de sustrato de la lipasa 2-17, que indican que el mejor sustrato para esta enzima son triacilglicerol con cadenas de aproximadamente 6 carbonos (Parra y col., 2007).

### **3.4.3 Determinación de variantes para la triada catalítica de la lipasa 2-17**

#### **3.4.3.1 Generación de Mutantes**

Para determinar la importancia de los aminoácidos seleccionados como candidatos de la triada catalítica, se generaron *in silico* variantes de la enzima para cada uno de estos aminoácidos. La serina y la histidina fueron reemplazadas por una asparagina, mientras que el aspartato fue reemplazado por una alanina, una asparagina y ácido glutámico. Las mutaciones generadas fueron denominadas: S239N, D361A, D361E, D361N y H391N.

#### **3.4.3.2 Análisis computacional de las mutaciones**

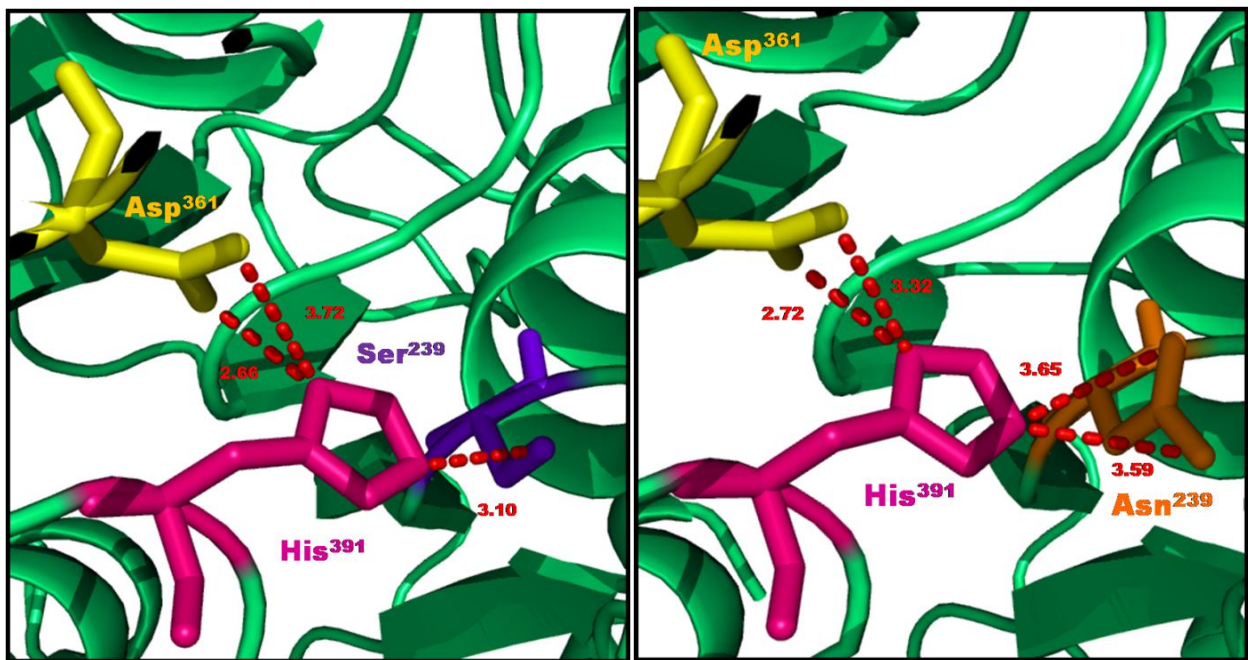
Mediante el modelamiento por homología de los variantes, se estudiaron los cambios producidos por las mutaciones a nivel estructural. Se realizó un análisis de la disposición espacial de los aminoácidos y su relación con los aminoácidos del sitio activo. Para ello mediante el programa Modeller se llevó a cabo el modelamiento de cada una de las mutaciones. Al igual que en la sección 2.3.1 el procedimiento para elaborar cada uno de los modelos consistió en comparar la secuencia de la lipasa mutada con la secuencia de proteínas cuyas estructuras fuesen obtenidas a través de cristalografía de rayos X o bien NMR. Se utilizó como molde para el modelo comparativo de cada mutación, la carboxilesterasa AFEST, la cual fue alineada con la secuencia de la lipasa 2-17 mutada para cada uno de los aminoácidos de la triada. Se generaron cerca de 20 modelos en cada caso y utilizando el índice DOPE se seleccionaron los mejores. Finalmente cada uno de los modelos fue validado mediante el servidor VADAR y refinado utilizando Modeller. Además se procedió a comparar los modelos obtenidos con modelos generados a través del servidor Swiss Model para asegurar la calidad y consistencia de cada uno de los modelos

(datos no mostrados). Una vez generados los modelos, éstos fueron analizados mediante Pymol y comparados con la enzima nativa.

- **Mutación S239N**

En el modelo de la mutación de la serina por la asparagina, no se observa catálisis porque si bien la asparagina cuenta con un oxígeno, el grupo amida no permite que este aminoácido se active como un buen nucleófilo.

Además es importante considerar que la asparagina es un aminoácido bastante más grande que a pesar de poder sustituir reactivamente a la serina, provoca un pequeño cambio en el posicionamiento de la triada que podría tener como consecuencia cambios en el posicionamiento del sustrato, lo que podría estar causando la abolición en su hidrólisis (Figura 34). Además las distancias atómicas entre los aminoácidos de la triada catalítica cambian, lo que podría afectar el correcto desplazamiento de cargas a través de estos aminoácidos.



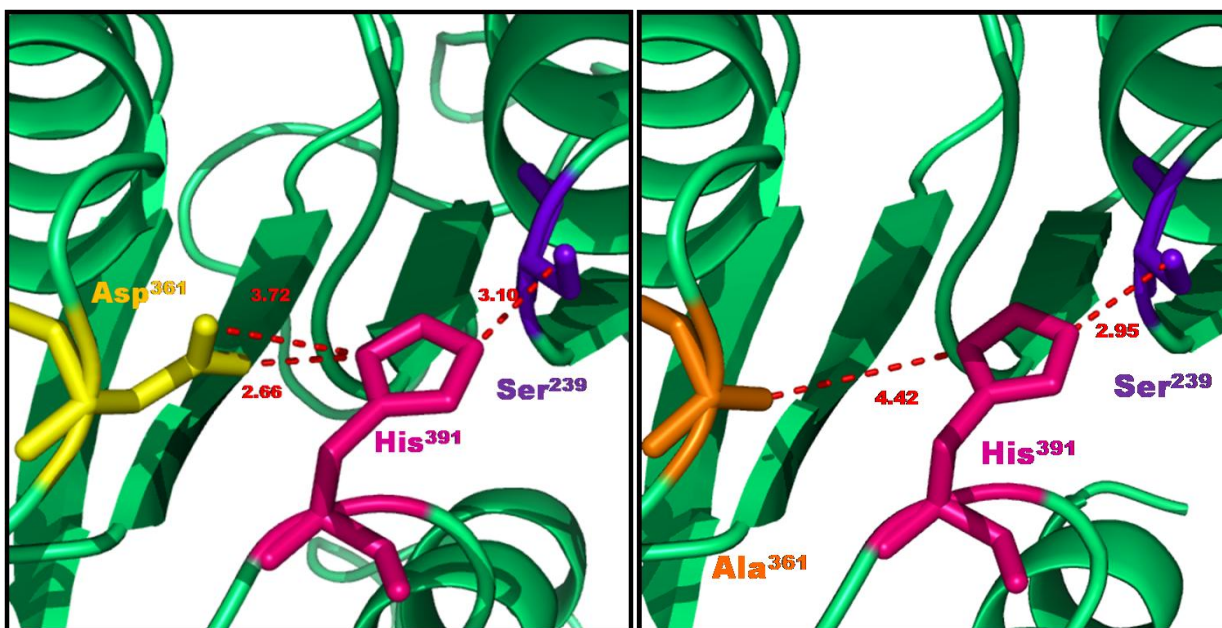
**Figura 34:** Triada catalítica de la lipasa 2-17 antes y después de la mutación S239N. **Izquierda:** Triada catalítica de la lipasa 2-17 con la serina 239. **Derecha:** triada catalítica de la lipasa 2-17 con la asparagina 239. Imágenes obtenidas con el programa Pymol.

- **Mutación D361A**

Para el caso del cambio del aspartato por una alanina, hay dos principales consecuencias que evitan la hidrólisis del sustrato. La primera corresponde a la lejanía entre los átomos de ambos

aminoácidos ya que la alanina al ser una molécula más pequeña no logra interactuar con la histidina, y la segunda y más importante debido a que el grupo radical de la alanina corresponde a un metilo y en el caso de que no hubiese un impedimento por la distancia de sus átomos, aún así no se podría formar enlace con el hidrógeno de la histidina dada la naturaleza no polar de la cadena lateral de la alanina (Figura 35).

El aspartato al ser un aminoácido ácido, cuyo oxígeno se encuentra cargado a pH fisiológico. Tiene la capacidad de unirse al hidrógeno de uno de los nitrógenos de la histidina, lo que permite hacer resonar las cargas, llevando a la activación del otro nitrógeno aromático. Este interactúa con el hidrógeno de la serina y por lo tanto permite el ataque nucleofílico de esta última sobre el carbono del glicerol.

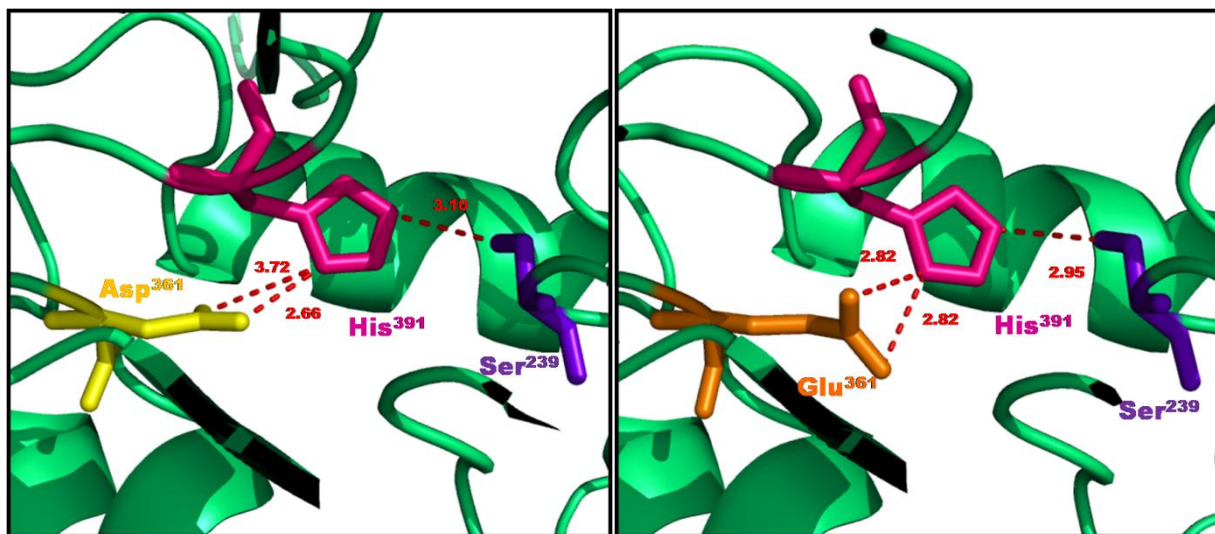


**Figura 35: Triada catalítica de la lipasa 2-17 antes y después de la mutación D361A. Izquierda:** Triada catalítica de la lipasa 2-17 con la aspartato 361. **Derecha:** triada catalítica de la lipasa 2-17 con la alanina 361. Imágenes obtenidas con el programa Pymol.

- **Mutación D361E**

Para la mutación de aspartato por glutamato, la situación es un poco diferente. El ácido glutámico tiene a diferencia del ácido aspártico, un carbono más en la cadena de su radical, por lo tanto, cuenta con una cadena lateral con carga negativa, que podría interactuar con el hidrógeno de la histidina y permitir la activación de su otro nitrógeno aromático. El problema fundamental en este caso, es que la presencia del carbono extra, hace que la cadena radical de este aminoácido tome otra orientación.

El cambio en la orientación del ácido glutámico tiene como consecuencia que las distancias atómicas entre el oxígeno reactivo del glutamato y el hidrógeno de la histidina cambien. Por su parte, la histidina se acerca a la serina, quedando a una distancia de 2.95 Å, pero cuya activación no se ve favorecida por el acercamiento de la histidina. Por un lado, enlaces muy débiles no favorecen la transferencia de cargas a través de la triada, y por otro, enlaces muy fuertes no permiten que la triada pueda recuperar su conformación inicial para completar el ciclo catalítico (Figura 36).

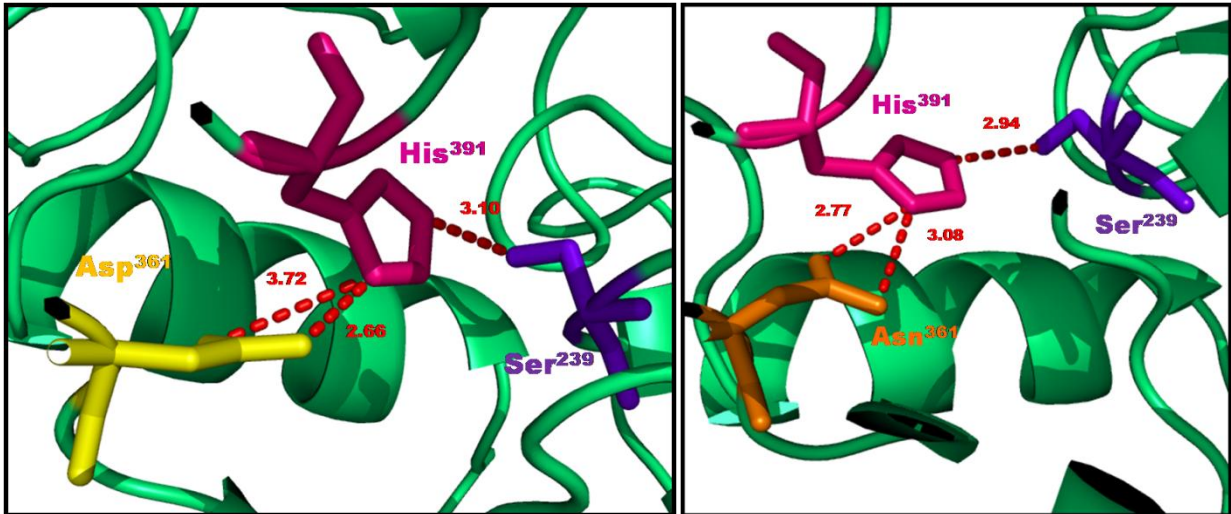


**Figura 36: Triada catalítica de la lipasa 2-17 antes y después de la mutación D361E. Izquierda:** Triada catalítica de la lipasa 2-17 con la aspartato 361. **Derecha:** triada catalítica de la lipasa 2-17 con el glutamato 361. Imágenes obtenidas con el programa Pymol.

- **Mutación D361N**

La asparagina es un aminoácido que se encuentra comúnmente al comienzo y al final de estructuras tipo  $\alpha$  hélice o bien en giros de hojas  $\beta$ . Esto es debido a su capacidad para formar enlaces con el esqueleto peptídico y estabilizar estructuras. Como se indica anteriormente, este aminoácido posee una cadena lateral con un grupo amida, que lo hace ser un aminoácido polar. Para el caso de la mutación del ácido aspártico por la asparagina, nuevamente es posible ver que las distancias atómicas entre los diferentes integrantes de la triada se ven afectadas (Figura 37), y por tanto el ordenamiento del sitio activo, debido a la orientación de la asparagina por su interacción con la histidina.

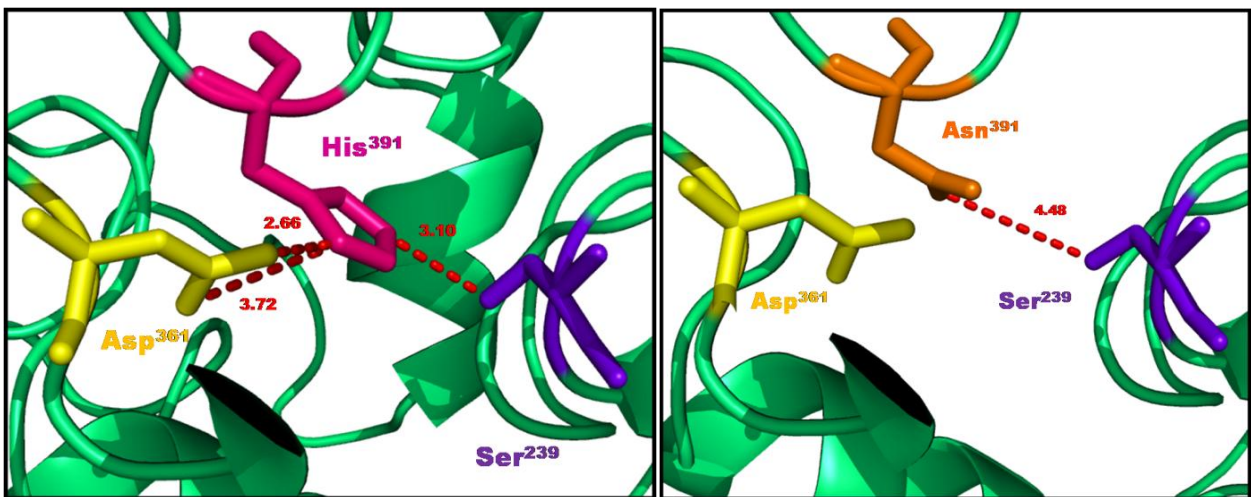




**Figura 37: Triada catalítica de la lipasa 2-17 antes y después de la mutación D361N.** Izquierda: Triada catalítica de la lipasa 2-17 con aspartato 361. Derecha: triada catalítica de la lipasa 2-17 con la asparagina 361. Imágenes obtenidas con el programa Pymol.

- **Mutación H391N**

En el cambio de la histidina por la asparagina es posible ver que la primera diferencia radica en la falta de interacción entre el aspartato y la asparagina. En el caso de la histidina, la interacción con el ácido aspártico hace que el grupo imidazol de la histidina adopte una orientación tal que permite su acción como base, aceptando al protón de la serina y formando un puente salino transitorio con el aspartato (Figura 38).



**Figura 38: Triada catalítica de la lipasa 2-17 antes y después de la mutación H391N.** Izquierda: Triada catalítica de la lipasa 2-17 con la histidina 391. Derecha: triada catalítica de la lipasa 2-17 con la asparagina 391. Imágenes obtenidas con el programa Pymol.

La principal función de la histidina en la catálisis del sustrato es activar a la serina para que realice el ataque nucleofílico al carbono del glicerol, haciendo resonar sus cargas a través del imidazol, y para ello las cargas deben estar correctamente distribuidas, y tanto las distancias y el posicionamiento de los átomos deben cumplir con las restricciones establecidas evolutivamente por estas enzimas. Por su parte, debido a que la asparagina no es capaz de aceptar protones ni de deslocalizar cargas, aunque la asparagina estuviese a la distancia correcta, no habría hidrólisis de sustrato.

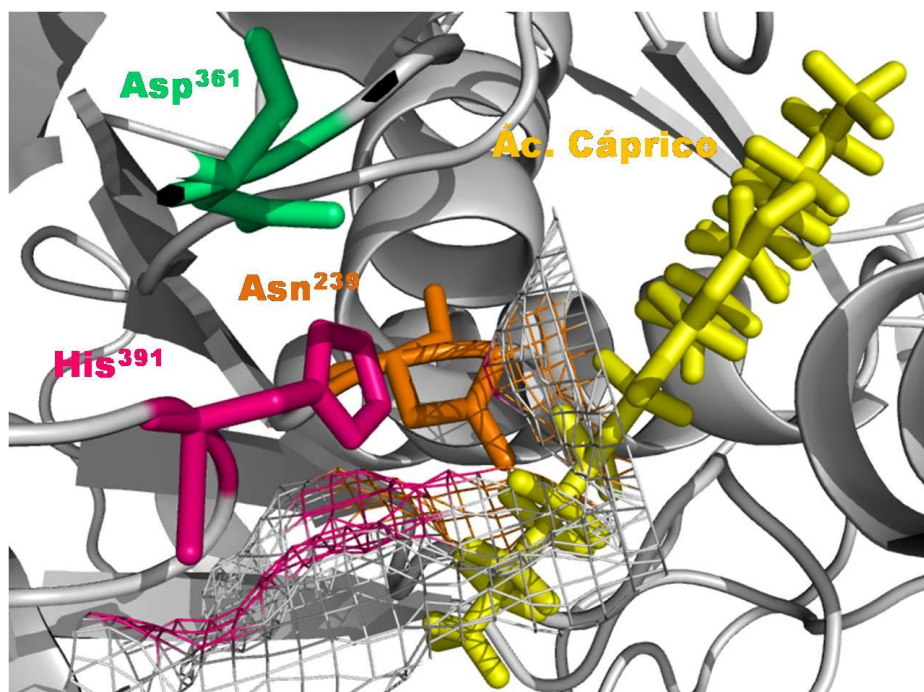
#### **3.4.1 Estudio de la interacción de mutantes lipasa-triacilglicerol mediante *Docking* Molecular**

Para conocer y comprender los resultados obtenidos para cada una de las mutaciones realizadas en la triada catalítica de la lipasa 2-17, se procedió a realizar *docking* molecular para cada una de ellas con el sustrato preferente por esta lipasa, triacilglicerol con cadenas de 6 carbonos.

Al igual que en la sección anterior, se realizaron dos tipos de *docking* molecular: *docking* ciego para permitir la unión del ligando en cualquier parte del receptor y posteriormente *docking* forzado para limitar el espacio de búsqueda a la triada mutada. Esto se repitió para cada una de las mutaciones simples. Para el *docking* ciego, la caja definida fue de 6Å x 35Å x 25Å (caja de 6 x 35 x 25, con puntos espaciados cada 1Å), al igual que se definió en la lipasa nativa, mientras que para el *docking* forzado se centró la búsqueda en los aminoácidos mutados. Se llevaron a cabo 5 simulaciones: S239N, D361A, D361E, D361N y H391N, correspondiente a cada una de las mutaciones.

En el caso del cambio de la serina 239 por la asparagina (Figura 39), no refleja una correcta hidrólisis del sustrato posiblemente debido a los problemas estéricos causados por la asparagina en el sitio activo y por lo tanto en la interacción con el sustrato.

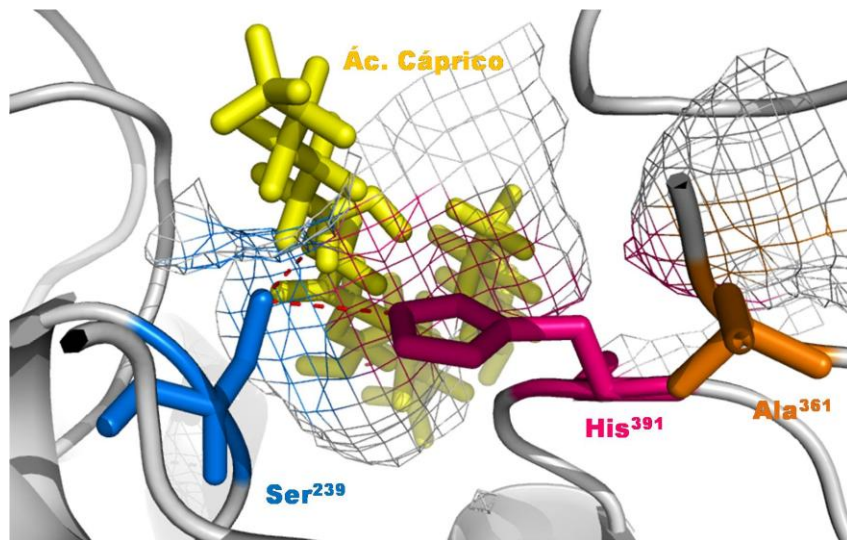
Las energías de interacciones entregadas por Autodock y Vina, fueron bastantes mayores a las obtenidas para la triada nativa reflejando los problemas en la interacción de la nueva triada con el sustrato (ver anexos 8.3.3). En el caso del análisis de la triada mutada sin sustrato, fue posible apreciar una interacción entre la histidina y la asparagina debido a la cercanía de ambos aminoácidos y un cambio en el posicionamiento de éstos. Sin embargo, al incorporar el sustrato, el cambio espacial de la histidina y de la asparagina fue notable, alejando a la histidina de la asparagina y por lo tanto es posible observar que no se forma una unión entre ellos, a pesar de la capacidad de ambos aminoácidos de establecer enlaces.



**Figura 39: Interacción mutante S239N y ácido caproico.** Triacilglicerol formado por ácidos caproico de 6 carbonos mostrado en amarillo. Triada catalítica mutada formada por asparagina 239, aspartato 361 e histidina 391 en anaranjado, verde y rosa respectivamente. Figura elaborada con el programa Pymol.

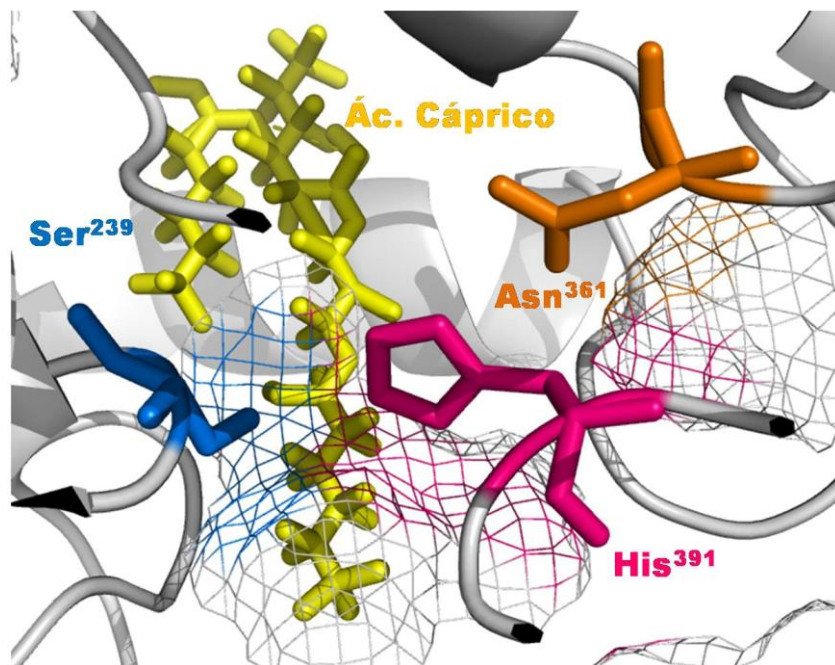
Para el caso de la mutación D361A en la cual el aspartato es reemplazado por una alanina, las diferencias entre las energías de interacción son aún mayores (ver anexos 8.3.3). Se crea una interacción entre el oxígeno de la serina y el oxígeno de uno de los ácidos grasos, pero cuya energía de interacción no permite la hidrólisis del sustrato (Figura 40). Particularmente, esta simulación presentó serias inconsistencias entre Vina y Autodock. Para ello, se realizó un *docking* molecular mediante el servidor Patch Dock (datos no mostrados), que realizó un modelo similar al obtenido por Vina entregando un soporte a la hipótesis que fundamenta los problemas de interacción por la naturaleza no reactiva de la cadena lateral de la alanina y la lejanía entre los átomos de la triada.

Al igual que en el caso del cambio de la serina por la asparagina, en teoría es posible suponer interacciones que permitan la catálisis del sustrato; sin embargo al hacer la simulación con el sustrato incorporado, es posible ver que no existen tales interacciones. Esto posiblemente se deba a que en presencia del aspartato, se forma un puente de hidrógeno entre el oxígeno reactivo de este aminoácido y el grupo imidazol adyacente presente en la histidina. Esta interacción es fundamental ya que orienta al grupo imidazol de tal manera que aumenta el pKa del grupo, formando un puente salino transitorio con el aspartato y permitiendo que al entrar en contacto con la serina, ésta se active y ataque al sustrato.



**Figura 40: Interacción mutante D361A y ácido caproico.** Triacilglicerol formado por ácidos caproico de 6 carbonos mostrado en amarillo. Triada catalítica mutada formada por serina 239, alanina 361 e histidina 391 en celeste, anaranjado y rosa respectivamente. Figura elaborada con el programa Pymol.

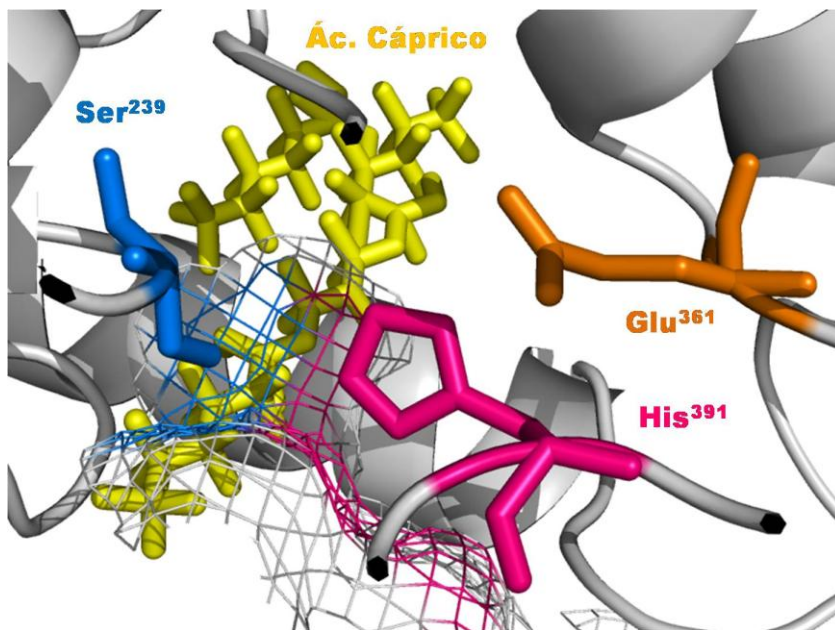
Para el caso de la asparagina, si bien es posible que el grupo amida interactúe con el nitrógeno no protonado del grupo imidazol, los resultados del *docking* molecular revelaron que en presencia del sustrato, la orientación de estos aminoácidos no favorece la activación de la serina (Figura 41).



**Figura 41: Interacción mutante D361N y ácido caproico.** Triacilglicerol formado por ácidos caproico de 6 carbonos mostrado en amarillo. Triada catalítica mutada formada por serina 239, asparagina 361 e histidina 391 en celeste, anaranjado y rosa respectivamente. Figura elaborada con el programa Pymol.

Al hacer el cambio con ácido glutámico, tal como se comentó en la sección 3.4.3.2, el carbono extra presente en la cadena lateral de este aminoácido orienta al oxígeno reactivo en otra dirección, impidiendo que éste interactúe con el hidrógeno del grupo imidazol de la histidina (Figura 42).

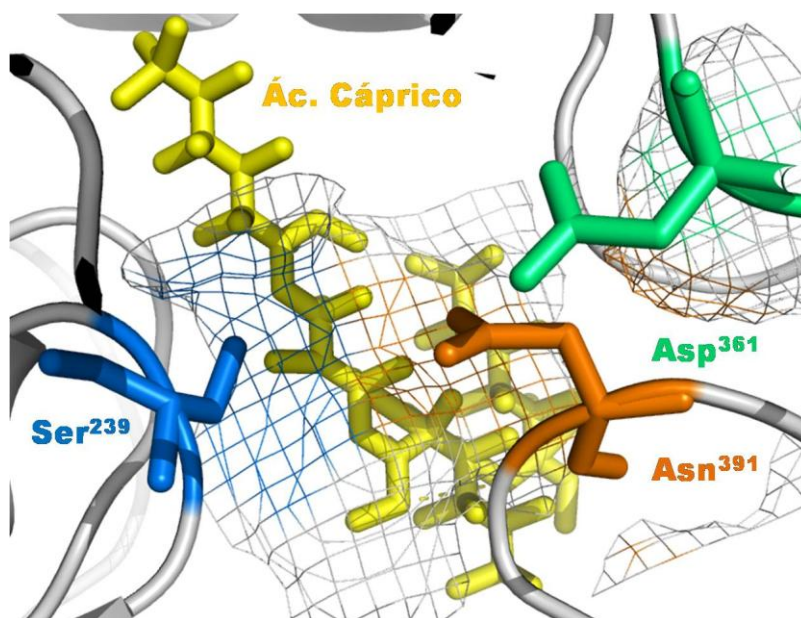
Así, el cambio en la orientación tiene como consecuencia que aumenten las distancias atómicas entre los elementos que interactúan para formar el puente de hidrógeno, y por lo tanto la no activación de la histidina conlleva a que la serina no se transforme en el poderoso nucleófilo que tiene la capacidad de atacar al carbono del grupo carbonilo del sustrato que permite la liberación del ácido graso.



**Figura 42: Interacción mutante D361E y ácido caproico.** Triacilglicerol formado por ácidos caproico de 6 carbonos mostrado en amarillo. Triada catalítica mutada formada por serina 239, glutamato 361 e histidina 391 en celeste, anaranjado y rosa respectivamente. Figura elaborada con el programa Pymol.

Finalmente, el *docking* molecular realizado para la mutación H391N (Figura 43) reveló que el cambio de la histidina por la asparagina no permite la hidrólisis del sustrato al no estar presente el grupo imidazol de la histidina que. En la enzima silvestre el imidazol es polarizado por el aspartato para actuar como base y aceptar un protón de la serina, resultando que ésta actúe como un nucleófilo mucho más fuerte. El acoplamiento del sustrato con la triada mutada mostró que la asparagina no es reactiva ni con el aspartato ni con la serina y que las energías de interacción son desfavorables para llevar a cabo la hidrólisis del sustrato (ver anexos 8.3.3).

Para todos los casos, los resultados obtenidos en el *docking* molecular para la hidrólisis del sustrato no fueron satisfactorios, indicando la importancia funcional de la triada catalítica que ha mantenido durante la evolución a estos tres aminoácidos, serina, aspartato e histidina como los protagonistas en gran parte del grupo de enzimas hidrolasas.



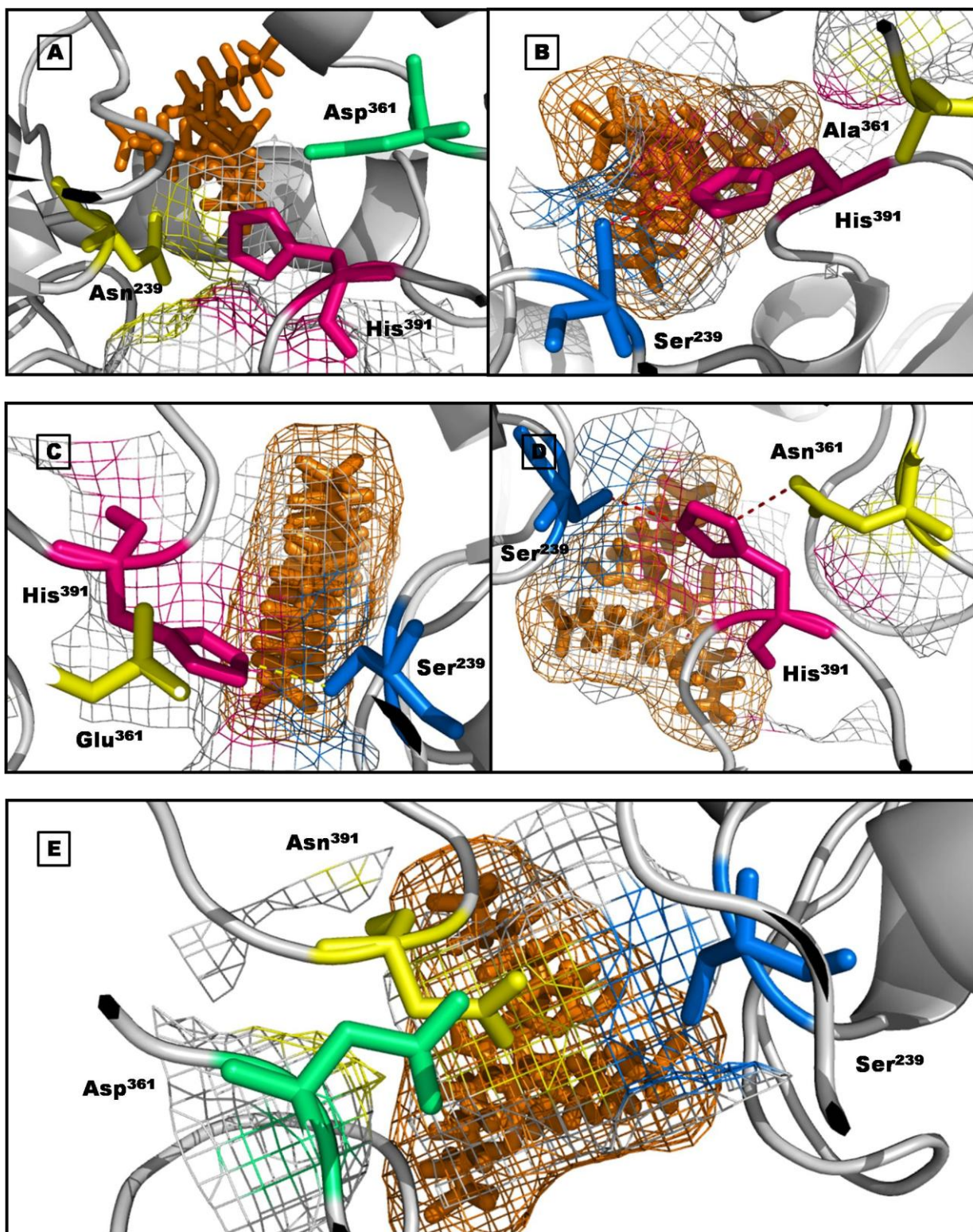
**Figura 43: Interacción mutante H391N y ácido caproico.** Triacilglicerol formado por ácidos caproico de 6 carbonos mostrado en amarillo. Triada catalítica mutada formada por serina 239, aspartato 361 y arginina 391 en celeste, verde y anaranjado respectivamente. Figura elaborada con el programa Pymol.

### 3.4.2 Estudio de la interacción de mutantes lipasa-triacilglicerol de cadenas largas, mediante *Docking Molecular*

Como se determinó en el *docking* molecular realizado para cada una de las mutaciones, uno de los principales problemas en la catálisis por los variantes fueron los cambios de las relaciones espaciales entre los aminoácidos involucrados, lo que llevó a la variación en las distancias atómicas entre estos aminoácidos y entre ellos con el sustrato. Para conocer su comportamiento en relación a este punto, se estudiaron los acoplamientos de cada una de las variantes con los sustratos analizados utilizados inicialmente para la lipasa nativa (triacilgliceroles con cadenas de 8, 10 y 16 carbonos).

Para el caso del triacilglicerol de 8 carbonos (ácido caprílico), en presencia de este sustrato se observó la interacción entre algunos aminoácidos, pero en todas las mutaciones las energías de afinidad no fueron suficientes para llevar a cabo la hidrólisis del sustrato (ver anexos 8.3.3) (Figura 44).

Para las mutaciones S239N y H391N no se observaron interacciones entre ninguno de los aminoácidos de la triada, ni entre alguno de ellos con el sustrato. Esto posiblemente se deba a que además de las diferencias espaciales que existen entre los átomos de cada aminoácido, la presencia del sustrato desordena aún más el sitio activo.



**Figura 44: Acoplamiento de mutantes de la lipasa 2-17 y triacilgliceroles con cadenas de 8 carbonos. A:** Mutación S239N. **B:** Mutación D361A. **C:** Mutación D361E. **D:** Mutación D361N. **E:** Mutación H391N. Para todos los casos la serina, el aspartato y la histidina están representadas con colores celeste, verde y rosa respectivamente. Todas las mutaciones están representadas en amarillo. El sustrato de 8 carbonos se encuentra en anaranjado.

Sin embargo, para las mutaciones D361A y D361E, es posible observar que la histidina efectivamente interactúa con la serina pero este puente de hidrógeno energéticamente no permite la activación de la serina para promover el ataque nucleofílico porque la histidina no está activada por el grupo ácido (ver anexos 8.3.3). Por otro lado, para el caso de la mutación D361N, es posible observar una interacción entre los aminoácidos de la triada. Sin embargo, el nitrógeno no protonado de la histidina forma un enlace de puente de hidrógeno con el grupo amida de la asparagina, que hace que al estar unida, adopte una conformación alternativa, la que conlleva a la interacción posterior con una molécula de agua, no permitiendo la hidrólisis del sustrato.

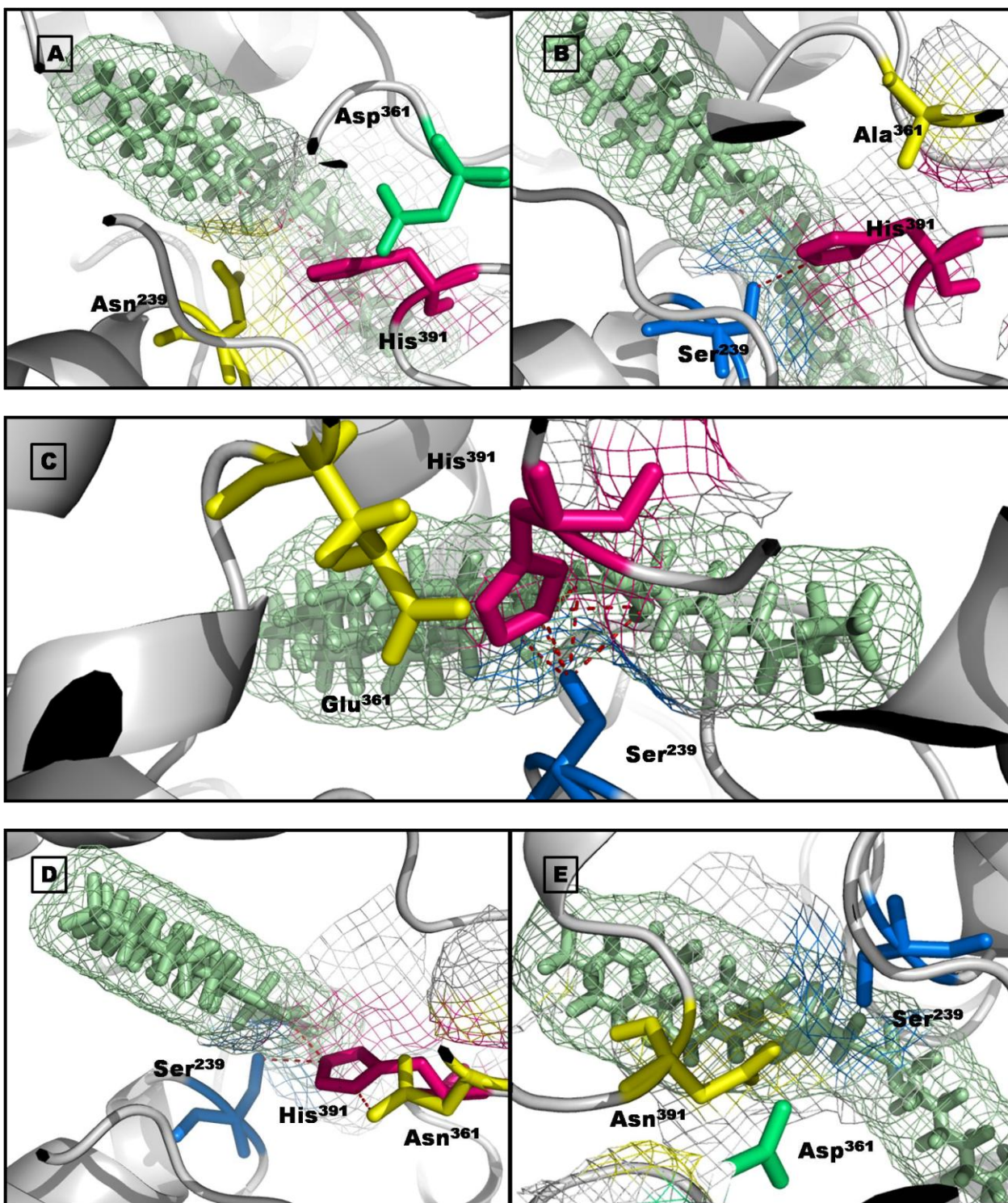
Para el caso del sustrato de 10 carbonos (Figura 45), la interacción de los aminoácidos en todos los escenarios es aun más inadecuada. Las mutantes S239N, D361E y D361N si bien presentan interacción de la triada con el sustrato, esta interacción es básicamente por la formación de puentes de hidrógeno con los oxígenos del sustrato, interacciones que no permiten la hidrólisis del grupo glicerol. Para la mutante D361A nuevamente es posible apreciar una interacción entre la histidina y la serina pero como se mencionó previamente, este puente de hidrógeno energéticamente no permite la activación de la serina (ver anexos 8.3.3). Por último, para la mutación H391N las energías entregadas por el docking son desfavorables y no se observa interacción alguna entre los participantes de la catálisis (ver anexos 8.3.3).

Finalmente, para el sustrato más grande, triacilglicerol de 16 carbonos (Figura 46), todas las mutaciones del ácido aspártico presentaron interacción entre la histidina y la serina pero no con el sustrato. Para las mutaciones S239N y H391N simplemente no hubo ningún tipo de interacción. En todos los casos se observó un reacomodo tanto en los aminoácidos de la triada, como en los aminoácidos aledaños al sitio activo. Como se comentó en la sección 3.4.2, las cadenas aciladas del triacilglicerol son tan grandes que provocan un leve cambio conformacional general en la proteína, ya que las hendiduras sn-1, sn-2 y sn-3 de la lipasa están formadas para acomodar cadenas mucho más pequeñas.

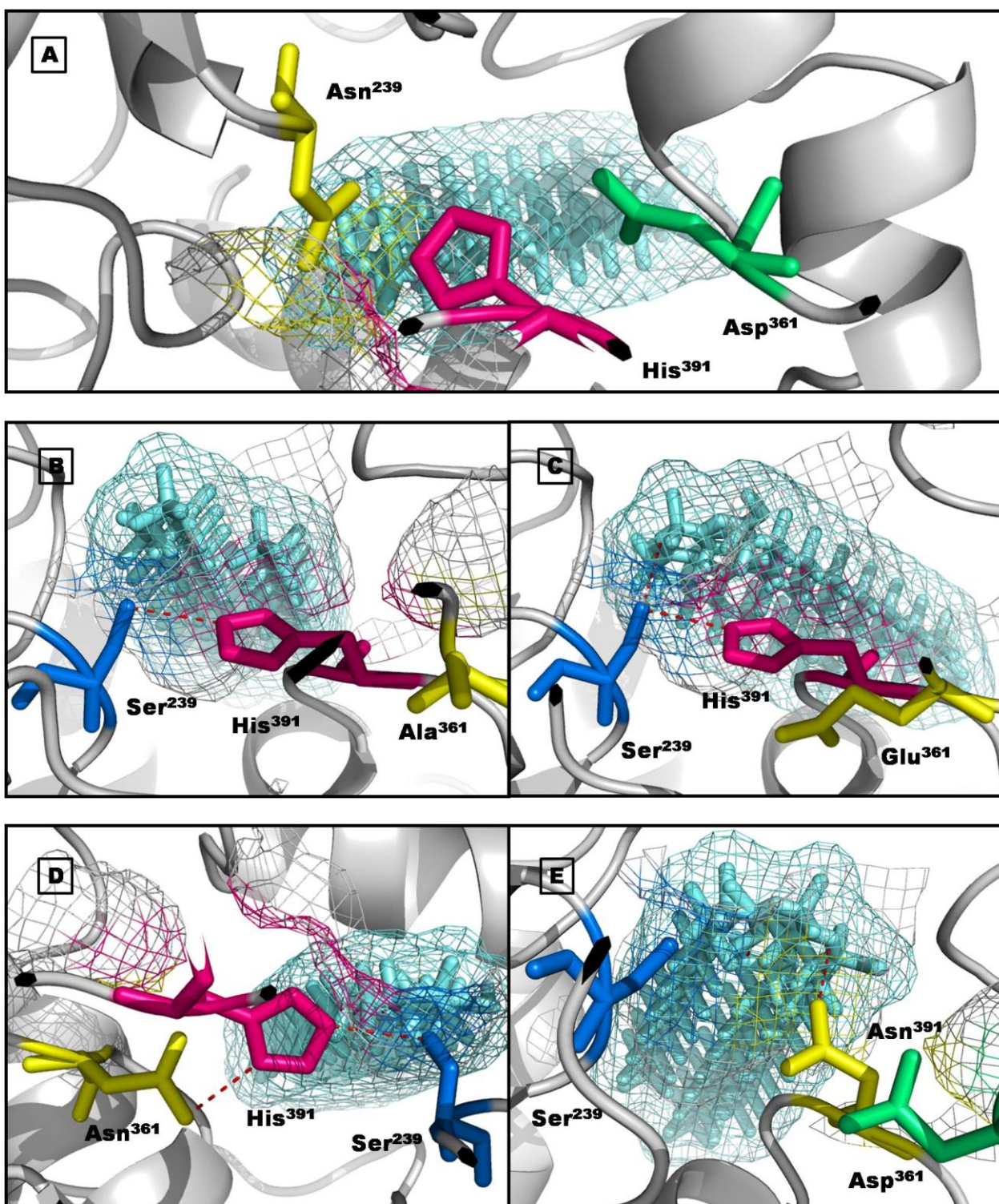
El acoplamiento con sustratos de dos carbonos no presentó resultados apreciables, ya que para ambos programas el posicionamiento de sustrato fue inespecífico tanto en el *docking* ciego como en el forzado (datos no mostrados).

Los resultados obtenidos para cada uno de los acoplamientos de los modelos mutantes con los diferentes tipos de triacilgliceroles, mostraron que a pesar del cambio en las distancias atómicas y los reacomodos de los aminoácidos en los diferentes casos, la especificidad por el sustrato no se vio alterada, y tanto los sustratos de cadena larga como los de cadena corta no exhiben un correcto modelo catalítico que permita justificar una mutación en la triada. Nuevamente, esto demuestra que la evolución ha optimizado la triada con estos aminoácidos como los protagonistas de estos sistemas catalíticos (Polgar, 2005).





**Figura 45: Acoplamiento de mutantes de la lipasa 2-17 y triacilgliceroles con cadenas de 10 carbonos.**  
**A:** Mutación S239N. **B:** Mutación D361A. **C:** Mutación D361E. **D:** Mutación D361N. **E:** Mutación H391N.  
 Para todos los casos la serina, el aspartato y la histidina están representadas con colores celeste, verde claro y rosa respectivamente. Todas las mutaciones están representadas en amarillo. El sustrato de 10 carbonos se encuentra en verde.



**Figura 46: Acoplamiento de mutantes de la lipasa 2-17 y triacilgliceroles con cadenas de 16 carbonos.** A: Mutación S239N. B: Mutación D361A. C: Mutación D361E. D: Mutación D361N. E: Mutación H391N. Para todos los casos la serina, el aspartato y la histidina están representadas con colores azul, verde claro y rosa respectivamente. Todas las mutaciones están representadas en amarillo. El sustrato de 16 carbonos se encuentra en celeste.

### 3.5 CONCLUSIONES

Gracias a los estudios de las funciones de las cadenas laterales de los aminoácidos en la catálisis serina-proteasa, se logró establecer que estas cadenas contribuyen a la mejora de la velocidad de la reacción favoreciendo la hidrólisis de los sustratos.

En términos generales, en proteasas de serina, cada uno de los tres residuos de la clásica triada catalítica tiene un rol específico en la generación del potencial nucleofílico. Al comparar las diferentes energías entregadas por el *docking* molecular realizado para la lipasa nativa y los diferentes sustratos, es posible apreciar que la menor energía y más favorable es la que presentó el sustrato de 6 carbonos, siendo de -6,9 Kcal/mol y la más desfavorable fue para el sustrato de 16 carbonos (-2,3 Kcal/mol). Esto posiblemente este dado por el enorme impedimento estérico que tiene el sustrato de mayor número de carbonos. Las cadenas aciladas, al no encajar correctamente en las hendiduras sn-1 sn-2 y sn-3, impiden que el ataque nucleofílico de la serina 239 sea óptimo para la hidrólisis del sustrato. Para los otros dos sustratos, las energías de afinidad fueron mayores a obtenidas para el sustrato de 6 carbonos.

En el caso de las mutaciones, las diferencias energéticas fueron aun mayores. Para el *docking* realizado con el sustrato de 6 carbonos, todas las variantes presentaron energías de afinidad mayores que la energía de la lipasa nativa, evidenciando una significativa diferencia entre las variantes y la lipasa nativa. Para los otros tres sustratos, las diferencias fueron menores debido a que las energías de actividad de la lipasa nativa eran desfavorables y mucho mayores que la del sustrato de 6 carbonos.

Existen otros factores importantes en el mecanismo catalítico como las condiciones del solvente, la temperatura y/o la presencia de cofactores, que influyen en los efectos de unión del sustrato y que permiten mantener los centros activos. Estos factores pueden causar ajustes estructurales transitorios que favorecen o perjudican la transferencia del protón del oxígeno de la serina, condicionando su nucleofilicidad y su capacidad de atacar al carbono del grupo carbonilo del sustrato.

Los resultados obtenidos *in silico* para la interacción de cada una de las enzimas mutantes con los diferentes sustratos, permiten dilucidar la importancia de estos aminoácidos para la actividad lipolítica de esta enzima, pero es necesario realizar estudios experimentales para validar estos resultados.

## **CAPÍTULO 4**

**Generación de variantes mediante Mutagénesis Sitio-Dirigida**

#### 4.1 RESUMEN

Como se mencionó en el capítulo anterior, el mecanismo catalítico de las lipasas se basa en un sistema de intercambio de cargas que consta de 4 etapas. Tras la unión del sustrato, se produce el ataque nucleofílico por parte del grupo hidroxilo de la serina catalítica sobre el enlace éster del lípido, lo que lleva a la rotura del enlace y a la formación de un intermediario entre el ácido graso y la serina nucleofílica. Seguidamente se libera el alcohol y se produce un segundo ataque nucleofílico por parte de una molécula de agua que ataca el enlace éster del intermediario transitorio, lo que produce la liberación del ácido graso y la regeneración del centro catalítico.

Utilizando los resultados obtenidos anteriormente para el modelo de la lipasa 2-17 y la identificación de los aminoácidos vinculados a la catálisis lipolítica, se diseñaron variantes de la enzima para cada uno de los aminoácidos candidatos. La serina y la histidina fueron reemplazadas por una asparagina, mientras que el aspartato, cuya principal función es activar a la triada, fue reemplazado por una alanina, una asparagina y ácido glutámico (mutaciones S239N, D361A, D361E, D361N y H391N).

Los genes de las enzimas mutantes fueron clonados en el vector de clonamiento pGEM-T Easy para luego ser incorporados al vector de expresión pMAL-c2E e integrados exitosamente en las cepas *E. coli* BL21(D3E). Los clones fueron inducidos con IPTG 0,3 mM para expresar cada una de las lipasas recombinantes mutadas, obteniéndose en cada caso una proteína de fusión soluble de aproximadamente 90 kDa, consistente con la suma de la lipasa de aproximadamente 47 kDa y de MBP de 43 kDa.

Cada una de las enzimas mutantes fue caracterizada y se determinaron sus curvas de progreso para el sustrato preferente, para diferentes sustratos, temperaturas y pHs. En todos los casos, los variantes de la enzima nativa no mostraron actividad catalítica indicando la importancia de éstos aminoácidos en la catálisis enzimática y respaldando a estos candidatos como la triada catalítica de la lipasa 2-17.

## 4.2 INTRODUCCIÓN

La capacidad catalítica de las enzimas reside en la organización espacial de los átomos activos, a través de los cuales se realizan todos los pasos químicos y estructurales de las reacciones enzimáticas. Las comparaciones estructurales y de secuencias de una gran variedad de proteínas han dejado en evidencia que los grupos catalíticos tienen diferentes patrones de variación, y que las proteínas por sí solas, presentan a veces una inesperada e incluso sorprendente relación estructural y evolutiva con otras familias de proteínas que tienen funciones completamente diferentes (Warshel, 1998; Pineda & Schwartz, 2006)

El primer sitio catalítico de una proteasa fue revelado hace más de 45 años, utilizando cristalografía de rayos X. Su análisis demostró inicialmente que dos residuos eran los principales encargados de la catálisis enzimática: Ser 195 e His57 (Matthews y col., 1967). Posteriores pruebas revelaron que ambos aminoácidos se encontraban suficientemente cerca como para formar un puente de hidrógeno entre ellos. Finalmente se identificó un aspartato como parte de estos aminoácidos catalíticos y se definió la triada Ser-His-Asp como la triada catalítica de las hidrolasas (Bruice & Benkovic, 2000). Si bien la triada catalítica fue observada por primera vez en proteasas de serina como Ser-His-Asp, existen una gran variedad de adaptaciones que han evolucionado de forma independiente (Dodson & Wlodawer, 1998). Las comparaciones entre los diferentes sistemas catalíticos que han llevado al reconocimiento de estas relaciones químicas y estructurales, fueron posibles gracias a la disponibilidad de resultados obtenidos desde análisis cristalográficos y gracias al desarrollo de nuevas herramientas y metodologías en ingeniería de proteínas.

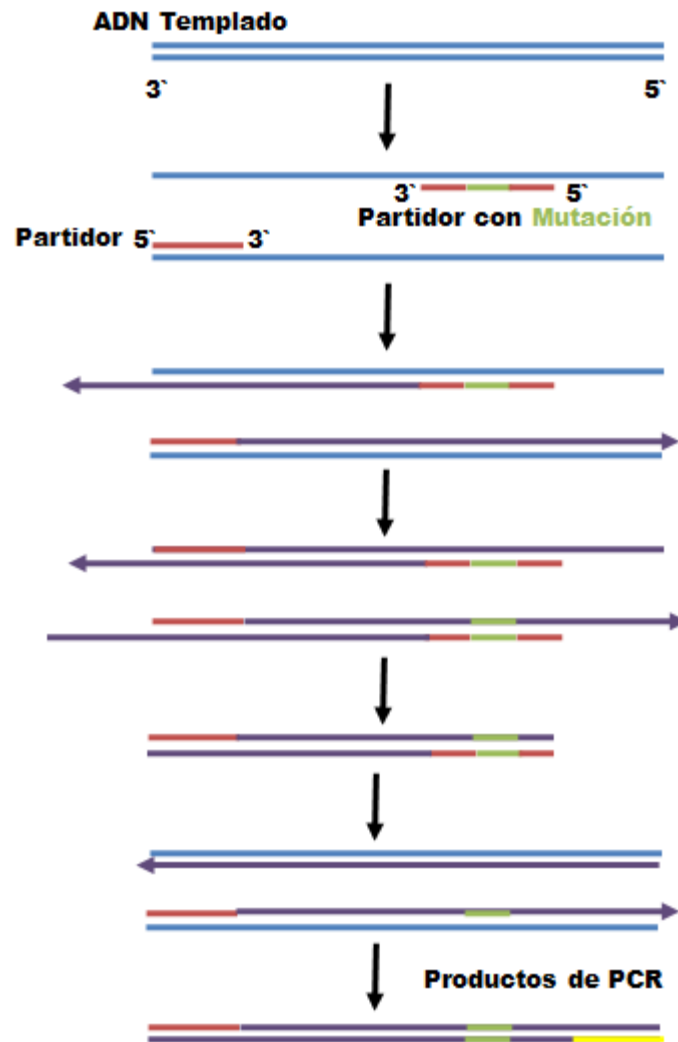
### 4.2.1 Mutagénesis Racional

El diseño racional de proteínas requiere tanto la disponibilidad de la estructura tridimensional de una enzima, como el conocimiento de la relación que existe entre la secuencia, la estructura y el mecanismo de funcionamiento. Si la información estructural de la enzima no está disponible, se puede utilizar la información de alguna estructura análoga cuya información sí se encuentre en las bases de datos (Bassegoda y col., 2012).

Toda esta información puede ser utilizada para identificar residuos específicos, que pueden ser modificados mediante mutagénesis (Sambrook & Russell, 2001), con el objetivo de mejorar una propiedad en particular como la especificidad, la termoestabilidad, la selectividad y la actividad entre otras. Los residuos seleccionados son entonces modificados mediante mutagénesis sitio-dirigida y las variantes obtenidas son expresadas, purificadas y analizadas en base a la propiedad buscada (Bassegoda y col., 2012).

#### 4.2.2 Mutagénesis Sitio Dirigida

Muchos temas relacionados a lipasas han sido discutido a lo largo de los últimos años, particularmente los relacionados con el sitio activo, actividad (incluyendo actividad específica, especificidad de sustrato, y especificidad por el largo de la cadena), estabilidad (incluyendo estabilidad térmica, a proteasas y oxidación), función de la tapa, unión de calcio, compatibilidad con detergentes y análisis de estructuras obtenidas por cristalografía de rayos X y dinámica molecular (Svendsen, 2000), y para ello, diversas técnicas han sido desarrolladas, con el objetivo de facilitar esta tarea y conocer más a fondo éste y otros tipos de proteínas.



**Figura 47: Esquema general Mutagénesis Sitio Dirigida.** ADN molde es copiado mediante partidores que incorporan la mutación, representados por una línea roja con un fragmento verde. Primer PCR se realiza con partidores de extremos (directo e inverso) y partidores con la mutación. Segundo PCR se realiza con los fragmentos mutados que actúan como macropartidores. Fragmentos finales completos con mutación incorporada.

En 1978 Michael Smith diseñó un método para cambiar un nucleótido específico en una secuencia de ADN con una alta eficiencia (Clyde y col., 1978). El primer paso consistió en sintetizar una secuencia complementaria a la secuencia que se deseaba mutar. Esta nueva secuencia era complementaria a la secuencia nativa excepto para el o los nucleótidos alterados. Este trabajo fue la base para lo que hoy en día se conoce como “Mutagénesis Sitio Dirigida” y ha sido la técnica precursora para gran parte de las modificaciones que se realizan para el mejoramiento y optimización de las propiedades de las proteínas.

Mediante mutagénesis sitio dirigida es posible modificar un aminoácido y evaluar su participación en la proteína, conocer sus propiedades a nivel catalítico, selectivo y de estabilidad, y estudiar modificaciones que podrían potenciar estas propiedades.

#### **4.2.3 Catálisis Enzimática**

Las enzimas adaptadas a bajas temperaturas exhiben estructuras comparables con sus homólogas mesófilas y termófilas pero se caracterizan por tener una eficiencia catalítica más alta a temperaturas más bajas y por su termoestabilidad baja a temperaturas moderadas (Zavodszky y col., 1998; Zaccai, 2000).

Para sobrevivir, los organismos psicrófilos necesitan enzimas capaces de catalizar eficientemente reacciones a temperaturas cercanas a los 0°C, temperatura a la que la mayoría de las otras especies no pueden crecer debido a su incapacidad para mantener sus flujos metabólicos adecuadamente (Zecchinon y col., 2001; Miyazaki y col., 2000). En comparación con sus homólogas mesófilas y termófilas, estas enzimas exhiben una mayor eficiencia catalítica a bajas temperaturas reflejada en un incremento en su número de recambio (aumento de  $k_{cat}$ ) y/o en el aumento de la afinidad por el sustrato (reducción de la  $K_M$ ) (Bae & Phillips, 2004; D’Amico y col., 2002).

En general se sabe que los determinantes estructurales de tales características cinéticas son responsables no sólo de estos nuevos óptimos de temperatura, sino también de la inactivación a moderadas y altas temperaturas (generalmente sobre 40 °C) para enzimas que trabajan en frío (Benkovic & Hammes-Schiffer, 2003; Eisenmesser y col., 2005). Se ha propuesto que la termolabilidad de enzimas activas a bajas temperaturas es una consecuencia de una mayor flexibilidad de su estructura molecular, en contraste con la termoestabilidad de las enzimas de organismos termófilos, que por el contrario, debe estar relacionada con la rigidez en sus cadenas peptídicas (Bae y col., 2004).

##### **4.2.3.1 Flexibilidad Enzimática**

La flexibilidad describe básicamente los cambios concertados que afectan los grados de libertad en una molécula, modificando la estructura completa sin destruirla (Dodson & Verma, 2006; Kumar y col., 2001). Se puede decir entonces, que una cadena polipeptídica es intrínsecamente flexible porque dentro



de sus esqueletos y sus cadenas laterales, gran parte de los enlaces covalentes tienen cierta libertad rotacional (Collins y col., 2003; Baldwin, 1986).

Kumar y col, (2001) propusieron que al hablar de flexibilidad en proteínas es importante considerar por un lado, los movimientos moleculares a gran escala, en los que porciones de la estructura se desplazan como cuerpos rígidos, movimiento relacionado a la unión de sustratos o regulaciones alostéricas (Halle, 2002). Este tipo de movimientos puede ser determinado comparando estructuras resueltas con y sin sustratos (o cofactores). Y por otro lado, se deben considerar los movimientos a pequeña escala, reflejados en un conjunto de isómeros conformacionales que constituyen el estado nativo de la proteína (Papaleo y col., 2006, 2007). Por lo tanto, dentro de las proteínas, podemos considerar movimientos de carácter local que comprenden pequeñas porciones de la proteína o movimientos de carácter global, que comprenden grandes porciones (Daniel y col, 2003; Agarwal, 2006; Böde y col., 2007).

Para mantener la funcionalidad de las moléculas, es necesario cubrir un amplio espacio conformacional, ya que deben ser capaces de adaptar toda su geometría para facilitar y permitir la unión de sustratos y en algunos casos de cofactores, poner en contacto las especies químicas reactivas en el complejo enzima-sustrato (Fields y col., 2001), crear un ambiente en el sitio activo con las características fisicoquímicas apropiadas para llevar a cabo la catálisis y liberar los productos resultantes de la reacción enzimática (Fleming & Richards, 2000; García-Viloca y col., 2004). Así, uno de principales temas de análisis evolutivo constituye la existencia de variaciones seleccionadas con el objetivo de incrementar la tasa catalítica en las enzimas (Bruice y Benkovic, 2000; D'Amico y col., 2003; Gromiha & Servaraj, 2004), particularmente en enzimas extremófilas (Bahar & Jernigan, 1997; Gomes & Steiner, 2004).

A bajas temperaturas, las enzimas comúnmente pierden gran parte de su actividad, y se comportan como elementos rígidos (Aghajari y col., 2003), pero los movimientos térmicos en estas condiciones básicamente son vibraciones armónicas, ya que a temperaturas suficientemente altas los átomos son capaces de superar las barreras energéticas que separan un estado conformacional de otro, y la estructura puede explorar un mayor número de distribuciones espaciales (Zaccai, 2000; Schiraldi & De Rosa, 2002).

Así también, el solvente ejerce una gran influencia sobre la flexibilidad de una proteína, y se ha observado que enzimas en medio acuoso pueden exhibir un aumento en la flexibilidad por el debilitamiento de las interacciones electrostáticas en la superficie de la molécula, debido a la alta constante dieléctrica del ambiente (Daniel y col., 2003; Alexandrov y col., 2005). Böde y col., (2007) propusieron que los puentes de hidrógeno que se forman entre la superficie de la molécula y el agua pueden inducir cambios en los niveles de energía de determinadas conformaciones espaciales, lo que

tiene como consecuencia un aumento en los posibles estados de la enzima, lo que podría tener como consecuencia, que la actividad de la enzima se vea afectada (Fenimore y col., 2002).

#### **4.2.3.2 Estabilidad Enzimática**

La estabilidad de las enzimas está relacionado con la diferencia energética entre el estado nativo y otros estados no nativos de la proteína (o proteína denaturada). Particularmente, la energía de estabilización neta de las proteínas es comúnmente muy baja (Zaccai, 2000) y corresponde a la energía de unos pocos puentes de hidrógeno (Jaenicke, 2000). Se ha propuesto que la inestabilidad de las enzimas, es decir la inducción a la denaturación, por la interrupción de algunas interacciones débiles en la estructura, podría estar relacionada con la flexibilidad y la actividad de la enzima.

La mantención de un balance adecuado entre estabilidad molecular y flexibilidad estructural es fundamental para la función de las proteínas. Así, la estabilidad es necesaria para evitar la denaturación de la proteína y asegurar la geometría molecular que permite la interacción con el sustrato, mientras que la flexibilidad es indispensable para la unión eficiente de sustratos, la catálisis y la liberación de los productos resultantes (Hoyoux y col., 2004; Henzler-Wildman y col., 2007). La necesidad de las enzimas por mantener un cierto umbral de flexibilidad que asegure su funcionalidad puede determinar los niveles de estabilidad que presentará la proteína (Dill & Bromberg, 2003). Así, la estabilidad global de una proteína no necesariamente responde a la flexibilidad de su sitio activo, y si se analiza la dinámica del sitio activo es posible establecer grandes diferencias con respecto al resto de la proteína (Karplus & Kuriyan, 2005; Rose y col., 2006).

Gran parte de las enzimas psicrófilas, presentan baja estabilidad estructural posiblemente debido al aumento en la flexibilidad local o global que permite disminuir la energía de activación de la reacción catalizada y por lo mismo aumentar la eficiencia catalítica (Hei & Clark, 1994; Svingor y col., 2001)

### 4.3 MATERIALES Y MÉTODOS

#### 4.3.1 Materiales

##### 4.3.1.1 Cepas bacterianas

Tabla V: Cepas bacterianas utilizadas

TIPO DE BACTERIA	GENOTIPO
<i>E. coli</i> cepa DH5 $\alpha$	F <sup>-</sup> $\Phi$ 80d <i>lac</i> $\Delta$ M15 $\Delta$ ( <i>lac</i> <i>aya-argF</i> ) U169 <i>recA1 endA1 hsdR178r<sub>k</sub><sup>-</sup>,m<sub>k</sub><sup>+</sup></i> ) <i>phoA supE44 <math>\lambda</math> thi-1 gyrA96 relA1</i> .
<i>E. coli</i> BL21(D3E)	F <sup>-</sup> <i>ompT hsdS<sub>B</sub> (r<sub>B</sub><sup>-</sup>m<sub>B</sub><sup>+</sup>) gal dcm</i> (DE3).

##### 4.3.1.2 Reactivos

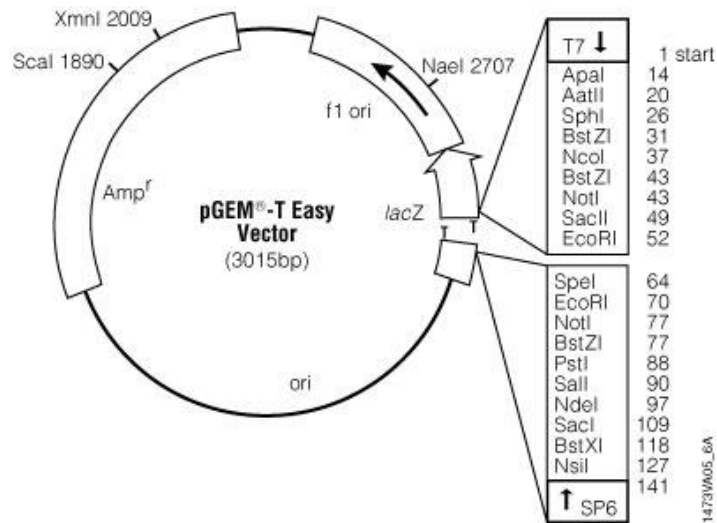
Tabla VI: Reactivos utilizados.

LABORATORIO	REACTIVOS
Promega (WI-USA)	Enzima de restricción <i>Hind</i> III. ADN <i>Taq</i> polimerasa, pGEM-T Easy.
Invitrogen (CA-USA)	T4 ADN ligasa, estándar de tamaño molecular 100pb y 1Kb, dNTPs, cepa <i>E. coli</i> DH5 $\alpha$ .
Sigma (MO-USA)	X-GAL, IPTG, goma arábica, <i>p</i> -NFC.
Thermo	Estándar de tamaño molecular preteñido de proteínas, enzima de restricción <i>Kpn</i> I.
Difco (DT- USA)	Triptona, extracto de levadura, medio LB preparado, agar.
New England Biolabs (MA-USA)	pMAL-c2E, resina de amilasa.
Novagen, Madison (WI-USA)	<i>E. coli</i> BL21(D3E).
Qiagen	Los sistemas de purificación de productos de PCR y de extracción de ADN desde geles de agarosa.
Winkler	Tris, SDS, glicerol.
J.T Backer	Cloruro de sodio.
Fermelo	Agarosa.
Merk	El resto de las sales, ácidos, bases, bromuro de etidio y solventes de grado analítico o de biología molecular.

### 4.3.1.3 Vectores de clonamiento y expresión

#### Vector de clonamiento pGEM-T Easy:

El vector de clonamiento pGEM-T Easy se utilizó para obtener un alto número de copias de fragmentos de ADN que se amplificaron previamente. El vector pGEM-T Easy posee una timina no apareada en el extremo 3' de cada una de las hebras en el sitio de inserción lo que permite la ligación de segmentos de ADN obtenidos por PCR, que tienen una adenina en el extremo 3' de cada hebra, debido a la actividad adenil-transferasa de la ADN *Taq* polimerasa (Figura 48).

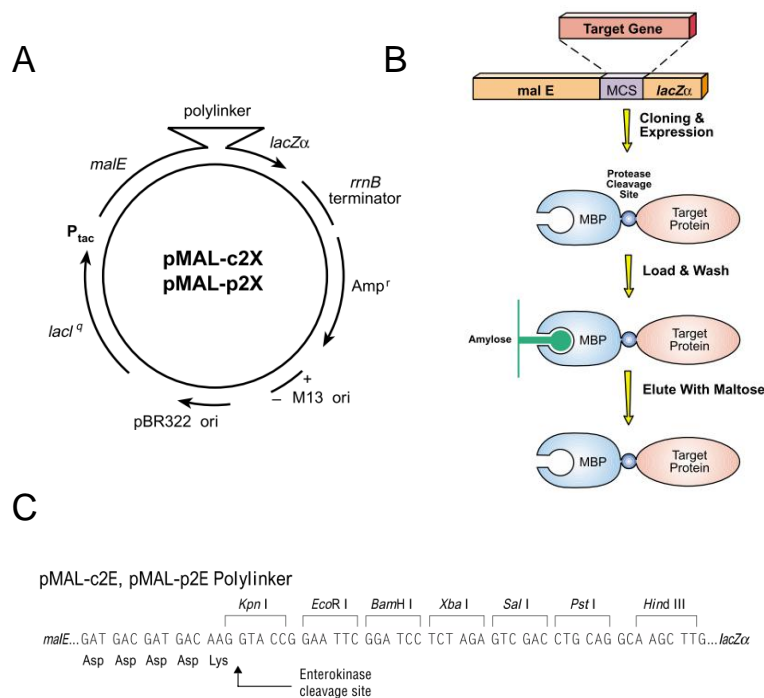


**Figura 48: Vector de clonamiento pGEM-T Easy.** F1 ori: ori de ADN de hebra simple derivado del bacteriófago F1. **Amp<sup>r</sup>**: gen de resistencia a ampicilina. **lacZ**: región codificante para la β-galactosidasa interrumpida por el sitio de múltiple clonamiento (Modificado de “Sistema de Vectores pGEM-T y pGEM-T Easy” de *Promega*, Cat. A1380).

#### Vector de expresión pMAL-c2E:

Para la expresión de las proteínas recombinantes en *E. coli* se utilizó el vector pMAL-c2E, en el cual el gen clonado se inserta en el sitio múltiple de clonamiento (MCS) río abajo del gen *malE*, que codifica para la proteína de unión a maltosa (MBP), resultando en la expresión de una proteína de fusión a MBP.

La fusión a MBP permite la estabilización de la proteína y su purificación mediante cromatografía de afinidad en la cual la proteína de fusión se une a maltosa. El vector pMAL-c2E tiene una delección en la secuencia señal *malE*, lo que resulta en la expresión citoplasmática de la proteína de fusión (Figura 49).



**Figura 49: Vector de expresión pMAL-c2E. (A)** Mapa del vector pMAL-c2X: **Ori pBR322**: ori bacteriano derivado de pBR322. **M13 ori**: Ori de ADN de hebra simple derivado del bacteriófago M13. **Amp<sup>r</sup>**: gen de resistencia a ampicilina. **LacI**: gen que codifica para represor lac. **P<sub>tac</sub>**: control del promotor híbrido tac. **LacZ**: región codificante para la β-galactosidasa interrumpida por el MCS. **malE**: secuencia codificante para la proteína MBP. **(B)** Esquema del clonamiento y expresión de la proteína de fusión. **malE**: secuencia codificante para la proteína de unión a maltosa (MBP). **(C)** Mapa del MCS del vector pMAL-c2E. La flecha indica el sitio de corte para la enzima enteroquinasa (Riggs y col., 1985, Manual “Proteína de fusión pMAL y sistema de purificación” de *New Englad BioLabs*, Cat. WO2007/ 120809).

#### 4.3.1.4 Medios de cultivo líquidos

##### Medio Luria-Bertani (LB)

El medio LB contiene por cada 1 L de solución: 10 g de triptona, 5 g de extracto de levadura y 10 g de cloruro de sodio. El volumen se llevó a 1 L con agua Milli-Q y el pH se ajustó a 7. Se esterilizó por autoclave y almacenó a temperatura ambiente.

##### Medio TB

Este medio contiene por cada 1 L de solución: 12 g de triptona, 24 g de extracto de levadura y 4 mL de glicerol. El volumen se llevó a 900 mL con agua Milli-Q, se esterilizó por autoclave y se agregó 100 mL de fosfato de potasio estéril 10 % pH 8,6.

#### **4.3.1.5 Medios de cultivo sólidos**

Para la preparación se agregó agar al 1,5 % p/v a los medios líquidos. Los antibióticos se esterilizaron por filtración antes de ser agregados al medio de cultivo.

#### **4.3.1.6 Soluciones Stock**

##### **IPTG 0,1 M**

Se preparó una solución stock de IPTG 0,1 M en agua Milli-Q estéril. Se esterilizó por filtración mediante un filtro con un tamaño de poro de 0,2  $\mu\text{m}$  (Sartorius, Hannover-Alemania). Se almacenó a -20  $^{\circ}\text{C}$ .

##### **XGal 50 mg/mL**

Se preparó una solución stock de 5-bromo-4-cloro-3-indolil-D-galactosido (XGal) a una concentración de 50 mg/mL en N,N'- dimetil-formamida. Se almacenó a -20  $^{\circ}\text{C}$ .

##### **Ampicilina 100 mg/mL**

Se preparó una solución stock de ampicilina a una concentración de 100 mg/mL en agua Milli-Q estéril. Se esterilizó por filtración utilizando un filtro estéril con un tamaño de poro de 0,2  $\mu\text{m}$ . Se almacenó a -20  $^{\circ}\text{C}$ .

#### **4.3.2 Metodología experimental**

##### **4.3.2.1 Cultivo de microorganismos**

###### **Crecimiento y mantención de microorganismos en medio sólido**

Las cepas de *E. coli* BL21 (DE3) y DH5 $\alpha$  stock se crecieron en placas con medio LB-agar a 37 $^{\circ}\text{C}$  durante 16 h y se mantuvieron a 4 $^{\circ}\text{C}$  por períodos menores a 30 días.

###### **Crecimiento de cepas en cultivos líquidos**

Las cepas de *E. coli* se inocularon en los medios correspondientes, suplementado con el antibiótico requerido. Se incubaron a 37  $^{\circ}\text{C}$  en un agitador orbital (MaxQ 4000, Barnstead, Lab-Line, Iowa, USA) a 200 rpm, hasta alcanzar la O.D<sub>600</sub> deseada.

##### **4.3.2.2 Amplificación de ADN a través de PCR**

Para la amplificación del gen de lipasa se utilizó la enzima KOD<sup>®</sup>, ya que genera menos errores durante la amplificación que la ADN Taq polimerasa. Las reacciones de PCR se realizaron en tubos *Eppendorf* de 600  $\mu\text{L}$  esterilizados por autoclave. El tamaño esperado del gen fue de aproximadamente

1300 pb. Para el PCR se mezclaron los siguientes componentes: solución amortiguador de PCR a una concentración final de 1X (Tris/HCl 200 mM, KCl 500 mM, pH 8,4), MgCl<sub>2</sub> 4 mM, dNTPs 10 mM (2,5 mM de cada dNTP), 10 pmoles de cada partidor (Tabla VII) y 1 unidad de *KOD hot start*. Se completó el volumen a 100 µL con agua Milli-Q estéril. Los tubos se sometieron al programa requerido, con una etapa inicial de denaturación a 95 °C por 5 minutos, seguida por 30 ciclos de amplificación que incluían 3 etapas: 1 min a 95 °C, 30 s a la temperatura dependiente de la T<sub>m</sub> de cada uno de los partidores utilizados y una etapa de extensión de 30 segundos a una temperatura de 72 °C. Finalmente se calentó 10 min a 72 °C. Los productos de PCR se analizaron mediante electroforesis en gel de agarosa 1 % p/v.

**Tabla VII: Partidores utilizados para la amplificación del gen de la lipasa 2-17 y para la amplificación de los mutantes.**

Nombre	Secuencia (5' → 3')	T <sub>m</sub> (°C)
LipFmut	GGTACCTATTTTACCAGTGCCATCGCTC	60,6
LipRMut	AAGCTTTCAGAGATTTGGCTCATCACCTC	61,8
LipS239NR	AAGCAACCACCTGCGTTATCACCAGAC	62,7
LipS239NF	ATAACGCAGGTGGTTGCTTGGCAGCAC	65,3
LipS239NC	TCGCGCATCGTATTGTCTGGTGATAAC	60,6
LipD361AR	TCATCTCGCAAGATGGCTAATTCCGCAAC	62,3
LipD361AF	TAGCCATCTTGCAGATGAAGGACTGG	61,7
LipD361AC	CAAGCTATGTCGTCGTTGCGGAATTAGCC	62,7
LipD361ER	TCATCTCGCAAGATCTCTAATTCCGCAAC	60,1
LipD361EF	TAGAGATCTTGCAGATGAAGGACTGG	59,3
LipD361EC	CAAGCTATGTCGTCGTTGCGGAATTAGAG	60,7
LipD361NR	TCATCTCGCAAGATGTTTAATTCCGCAAC	59,5
LipD361NF	TAAACATCTTGCAGATGAAGGACTGG	58,6
LipD361NC	CAAGCTATGTCGTCGTTGCGGAATTAAC	59,9
LipH391NR	CAAGTTAATAAAACCGTTCGGTGCGCCTAG	61,1
LipH391NF	TAAACATCTTGCAGATGAAGGACTGG	58,6
LipD391NC	CTTATACAGTGCTAGGCGCACCGAAC	61,1

#### **4.3.2.3 Electroforesis de ADN**

Las electroforesis de ADN se realizaron en una cámara horizontal Horizon-58 (Gibco-BRL Life Technologies, Inc., MA-USA). Cada muestra se mezcló en proporción 5:1 con solución amortiguador de carga 6X (glicerol 30% p/v y azul de bromofenol 0,25% p/v) y luego se cargaron en geles de agarosa de una concentración ajustada al tamaño del fragmento de ADN. Éstos se hicieron a concentraciones de 1,5 % p/v para fragmentos de 1000 pb. La agarosa se preparó en solución amortiguadora TAE 1X: Tris-Acetato-EDTA (Tris 40 mM, acetato 20 mM y EDTA 1 mM) y bromuro de etidio 0,2 µg/mL. Se incluyó estándar de peso molecular de tamaño 1 Kpb y 100 pb. Las electroforesis se realizaron bajo voltaje constante de 100 V (fuente de poder Model 500 de BRL Life Technologies, Inc., MD-USA). Se utilizó como solución amortiguador de corrida TAE 1X. Finalmente se observó el gel sobre un transiluminador UV y se fotografió.

#### **4.3.2.4 Purificación de ADN a partir de geles de agarosa**

Con el fin de purificar fragmentos de ADN desde geles de agarosa, se cortó la banda de ADN desde gel, se colocó en un tubo *Eppendorf* de 1,5 mL esterilizado por autoclave, y luego se utilizó el sistema QIAEX II Gel Extraction, de acuerdo a las instrucciones del fabricante. Finalmente, el ADN se eluyó en agua Milli-Q estéril a pH 8, previamente calentada a 50 °C.

#### **4.3.2.5 Ligación de fragmentos de ADN a vectores de clonamiento y expresión**

##### **Ligación en el vector de clonamiento pGEM-T Easy**

Para la ligación se utilizó como inserto los productos de PCR purificados desde geles de agarosa. La mezcla de reacción se hizo en un volumen final de 10 µL y contenía solución amortiguador de ligación 2X, 50 ng de vector pGEM-T Easy, 3,5 µL de inserto y 1 unidad de enzima T4-ADN ligasa. Se dejó ligando por 16 h a 4 °C. El producto de ligación se utilizó para transformar células *E. coli* DH5α electrocompetentes.

##### **Ligación en el vector de expresión pMAL-c2E**

La reacción de ligación se realizó en un volumen final de 10 µL. Consistió en una mezcla que contenía 100 ng de vector digerido con las enzimas KpnI y HindIII, 300 ng de inserto digerido con las enzimas de restricción KpnI y HindIII, solución amortiguador T4 ADN ligasa 5X (Tris/HCl 250 mM pH 7,6, MgCl<sub>2</sub> 50 mM, ATP 5 mM, DTT 5 mM, polietilenglicol-8000 25 % p/v) y 1 µL de T4 ADN ligasa. La mezcla de reacción se dejó a 16 °C por 16 h.



#### 4.3.2.6 Transformación de células electrocompetentes

Para la transformación de células electrocompetentes *E. coli* BL21 (DE3) o DH5 $\alpha$  se agregó 1  $\mu$ L de mezcla de ligación a 20  $\mu$ L de células. La electroporación se realizó en un equipo Cell-Porator<sup>®</sup> Electroporation System (Gibco-BRL Life Technologies, Inc., MD-USA) en las siguientes condiciones: 420V, 330 $\mu$ F, baja impedancia, una tasa de carga rápida y un tiempo máximo de 2,5 mseg. Las células transformadas se mezclaron con 1 mL de medio LB y se incubaron a 37  $^{\circ}$ C durante 1 h y con agitación de 200 rpm. Luego se centrifugó a 13.000 rpm (centrífuga Eppendorf 5403 de Eppendorf AG, Hamburgo, Alemania) durante 1 min y se descartaron 900  $\mu$ L del sobrenadante. El precipitado se diluyó en el sobrenadante remanente (100  $\mu$ L) y se plaqueó sobre placas de agar LB estériles que contenían ampicilina 100  $\mu$ g/mL, X-gal 50  $\mu$ g/mL e IPTG 1 mM. Las placas de transformación se incubaron a 37 $^{\circ}$ C durante 16 h. Para su mantención se conservaron a 4  $^{\circ}$ C.

#### 4.3.2.7 PCR de colonias

El PCR de colonias se realizó para corroborar la presencia de inserto en las colonias blancas seleccionadas como clones positivos. Las colonias blancas seleccionadas se transfirieron mediante una punta estéril desde la placa de transformación a las placas cuadrículadas y numeradas que contenían agar LB agar suplementado con ampicilina, IPTG y X-gal y a un tubo Eppendorf con 100  $\mu$ L de agua Milli-Q estéril. La suspensión (tubo Eppendorf) se incubó a 100  $^{\circ}$ C durante 10 min y se tomaron 5  $\mu$ L, los que se agregaron sobre 15  $\mu$ L de mezcla de PCR. La presencia de inserto se confirmó utilizando los partidores pMALLipF y pMALLipR, para el cual se utilizó el protocolo de PCR que se describe a continuación:

Las reacciones de PCR se realizaron en tubos *Eppendorf* de 600  $\mu$ L esterilizados por autoclave. Para el PCR se mezclaron los siguientes componentes: solución amortiguador de PCR a una concentración final de 1X (Tris/HCl 200 mM, KCl 500 mM, pH 8,4), MgCl<sub>2</sub> 4 mM, dNTPs 10 mM (2,5 mM de cada dNTP), 10 pmoles de cada partidor y 1 unidad de ADN Taq polimerasa. Se completó el volumen a 15  $\mu$ L con agua Milli-Q estéril. Los tubos se incubaron en un termociclador (modelo PTC-100 de MJ Research, Inc., MA-USA) y se sometieron al programa requerido, que incluía una etapa inicial de denaturación a 94  $^{\circ}$ C por 5 minutos, seguida por 30 ciclos de amplificación que incluían 3 etapas: 30 s a 94  $^{\circ}$ C, 30 s a 60  $^{\circ}$ C dada la T<sub>m</sub> de los partidores utilizados y una etapa de extensión de un minuto por cada kilobase de ADN que se quiera amplificar, lo que se hace a una temperatura de 72  $^{\circ}$ C, que corresponde a la temperatura óptima para la función catalítica de la Taq ADN polimerasa, que para este caso fue de 30 s. Finalmente se calentó 5 min a 72  $^{\circ}$ C. Los productos de PCR de colonias se analizaron mediante electroforesis en gel de agarosa 1 % p/v.

#### 4.3.2.8 Minipreparación de ADN plasmidial

El ADN plasmidial se extrajo a partir de un cultivo en 4 mL de medio LB incubado por 16 h a 37° C y 200 rpm. El cultivo se centrifugó a 13.000 x  $g$  (Centrífuga *Eppendorf* 5403 de Eppendorf AG, Hamburgo, Alemania) por 1 min en un tubo Eppendorf de 1,5 mL previamente esterilizado por autoclave. El proceso se repitió hasta completar los 4 mL necesarios para obtener un precipitado de bacterias. A continuación, se utilizó el sistema QIAprep Spin Miniprep Kit. La extracción se realizó según las instrucciones del fabricante.

#### 4.3.2.9 Digestión de ADN plasmidial

Las enzimas *KpnI* y *HindIII* se utilizaron para digerir el vector pGEM T-Easy con inserto para escindir el gen de lipasa ligado en él y el vector pMAL-c2E para su linealización. La reacción de digestión final fue de 20  $\mu$ L y contenía 4  $\mu$ L de solución amortiguador 5X, 1  $\mu$ L de cada enzima y 100 ng de vector para el caso de pGEM T-Easy con inserto y 500 ng del vector de expresión pMAL-c2E para el caso de la linealización. La mezcla se llevó al volumen final con agua Milli-Q estéril. Para las digestiones dobles se utilizó la solución amortiguador de la enzima *HindIII*, recomendada por *New England Biolabs* y 1  $\mu$ L de cada enzima. Se dejó reaccionando durante 2 h a 37 °C y el resultado se analizó mediante un gel de agarosa.

#### 4.3.2.10 Inducción de la síntesis de proteínas recombinantes mediante el sistema de expresión *E. coli*/pMALc-2E

La inducción de la expresión del gen se realizó según lo recomendado por el manual de “Proteína de fusión pMAL y sistema de purificación” de *New England BioLabs* con algunas modificaciones. Se inocularon 20 mL de medio LB/TB y ampicilina 100  $\mu$ g/mL con una colonia de células BL21 (DE3) recombinantes con el vector pMALc-2E que tenía incorporado el inserto. El inóculo se dejó creciendo a 37 °C con agitación de 200 rpm por 16 h. Al día siguiente se midió la densidad óptica a 620 nm ( $OD_{620}$ ) y se determinó el volumen necesario para inocular 250 mL de medios LB/TB y ampicilina con el cultivo de noche para una  $OD_{620}$  inicial de 0,05 nm. El cultivo se creció con agitación de 200 rpm a 37 °C hasta alcanzar una  $OD_{620}$  de 0,6 en aproximadamente 2 h, se tomó una muestra de 1 mL, se centrifugó (centrífuga Eppendorf 5403 de Eppendorf AG, Hamburgo, Alemania) y el precipitado se resuspendió en 100  $\mu$ L de solución amortiguador SDS-PAGE 5X para luego guardarlo a -20 °C (*muestra1: células sin inducir*). Al resto del cultivo de células se le agregó IPTG para una concentración final de 0,3 mM. Se continuó la incubación a 18 °C con agitación de 200 rpm. Pasadas 4 h se tomó una muestra de 1 mL, se

centrifugó y el precipitado se resuspendió en 100  $\mu$ L de solución amortiguador SDS-PAGE 5X para después guardarlo a -20 °C (*muestra 2: células inducidas*).

#### **4.3.2.11 Purificación de la proteína recombinante mediante el sistema de expresión *E. coli* BL21 (DE3)/pMALc2E**

Las células que expresaban la lipasa se recolectaron mediante centrifugación a  $4.000 \times g$  por 10 min (Centrífuga Sorvall® RC-28S y rotor GS-3 Sorvall®, DuPont, CT, USA). El sobrenadante se descartó y el precipitado se suspendió en 30 mL de solución amortiguador de columna (Tris/HCl 20 mM, NaCl 200 mM, EDTA 1 mM). Las células se congelaron a -20 °C durante 16 h. Luego se descongelaron en agua fría y se sonicaron con 10 pulsos de 30 s. Se centrifugó a  $9.000 \times g$  y 4 °C durante 20 min (Centrífuga Sorvall® RC-28S y rotor GS-3 Sorvall®, DuPont, CT, USA). El precipitado correspondió a material insoluble. Del sobrenadante, que correspondió al extracto crudo, se tomó una muestra de 5  $\mu$ L y se agregó solución amortiguador SDS-PAGE 5X. Todo esto se guardó a -20 °C (*muestra 3: extracto crudo*). El resto del sobrenadante se diluyó 5 veces en amortiguador de columna y se guardó para ser purificado por cromatografía, pero antes se hizo una prueba a escala piloto. Se agregaron 200  $\mu$ L de resina de amilasa en un tubo *Eppendorf* y se centrifugó a  $13.000 \times g$  durante 1 min (Centrífuga *Eppendorf* 5403 de *Eppendorf* AG, Hamburgo, Alemania). Luego se descartó el sobrenadante y la resina se diluyó en 1,5 mL de solución amortiguador de columna. Se centrifugó en las mismas condiciones anteriores 2 veces. Se mezclaron 50  $\mu$ L de resina con 50  $\mu$ L de extracto crudo y se incubó en hielo durante 15 min. Se centrifugó a  $13.000 \times g$  durante 1 min y el precipitado se lavó con 1 mL de solución amortiguador de columna, el que se volvió a centrifugar en iguales condiciones para finalmente suspender la resina en 50  $\mu$ L de solución amortiguador SDS-PAGE 5X (*muestra 4: Proteína de fusión unida a amilosa*). Con el objetivo de purificar la lipasa quimérica, se cargó el resto del volumen del extracto crudo en una columna PD10 (Bio-Rad Laboratories, CA-USA) empacada con 2 mL de resina de amilosa. Para esto se lavó dos veces con 5 mL de solución amortiguador de columna. Luego se pasó el extracto completo que correspondía a 150 mL aproximadamente. La columna se lavó con 5 mL de solución amortiguador de columna. El lavado se repitió dos veces y se procedió a eluir la proteína con 4 mL de solución amortiguador de columna más maltosa 10 mM. La columna de amilosa se regeneró con la siguiente secuencia de lavados: 6 mL de agua destilada, 6 mL de SDS 0,1 % p/v, 2 mL de agua destilada y 6 mL de solución amortiguador de columna. La columna se guardó en etanol 20 % v/v.

#### **4.3.2.12 Electroforesis en geles de poliacrilamida en condiciones denaturantes (SDS-PAGE)**

Las separaciones electroforéticas en condiciones desnaturalantes se realizaron en una cámara vertical MiniProtean II (Bio-Rad Laboratories Inc., CA-USA). Las muestras se mezclaron en proporción 4:1 con solución amortiguador de carga 5X (Tris/HCl 60 mM, glicerol 25 % v/v, SDS 2 % p/v, 2-mercaptoetanol 14,4 mM y azul de bromofenol 0,1 % p/v pH 6,8), se denaturaron por ebullición durante 5 min y posteriormente se cargaron en geles de acrilamida. Las electroforesis se realizaron bajo voltaje constante de 200 V (fuentes de poder Power Pac 1000 de Bio-Rad Laboratories Inc., CA-USA o Power Supply EPS 3500XL de Pharmacia Biotech AB, Uppsala, Suecia). Como solución amortiguador de corrida se utilizó Tris-glicina-SDS (Tris 25 mM, glicina 192 mM y SDS 0,1 % p/v) (Bollag y col., 1996).

#### **4.3.2.13 Medición de la actividad lipolítica de la proteína mediante ensayo en medio líquido**

Para el ensayo estándar se preparó como sustrato una solución stock de *p*-nitrofenil caprato 0,1 M en isopropanol (Diez veces el valor de  $K_m$  de la lipasa 2-17), y una solución amortiguadora de ensayo Tris/HCl 20 mM pH 8, que contiene Triton X-100 4,4 mg/mL y goma arábica 1,1 mg/mL. La solución sustrato/amortiguador se preparó al momento de realizar el ensayo en una proporción 1:35 (solución stock de sustrato: solución amortiguador de ensayo) para un volumen final de 900  $\mu$ L. La solución amortiguadora de ensayo se calentó a 45°C y se agregó el sustrato lentamente bajo agitación continua. Se dejó enfriar 2 min y se utilizó en el ensayo, para lo cual se mezcló esta solución con 100  $\mu$ L de enzima. En esta reacción enzimática, el sustrato *p*-nitrofenil caprato es hidrolizado por la lipasa formando el producto *p*-nitrofenol de color amarillo. La reacción se siguió espectrofotométricamente a 410 nm, longitud de onda a la cual se mide la formación de producto. Se usó como referencia la solución amortiguadora de ensayo. El blanco se preparó mezclando la solución sustrato/solución amortiguador con la solución amortiguadora en que se encontraba la proteína, en el mismo volumen que la muestra (hidrólisis no enzimática).

#### **4.3.2.14 Determinación de la concentración de proteínas**

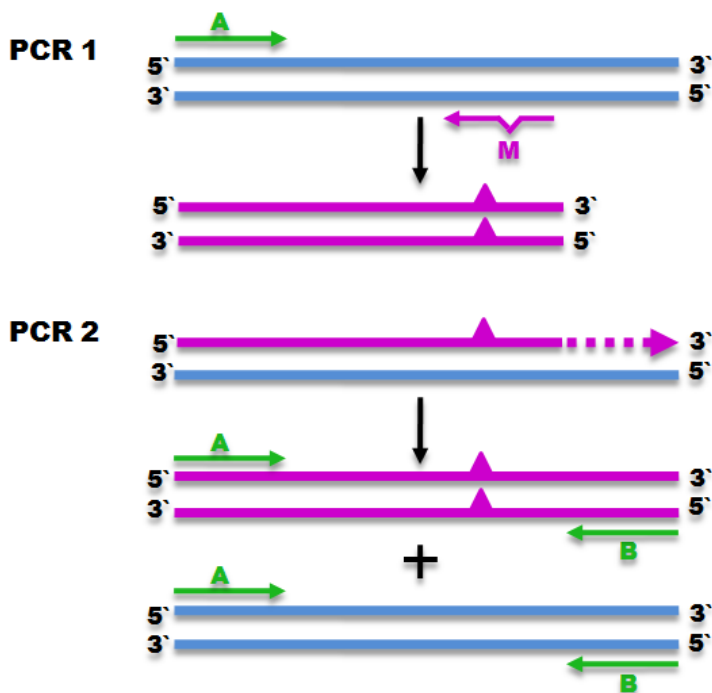
La determinación de la concentración de proteína total una vez purificada se realizó mediante el método de Bradford. Para ello, se mezcló: 2,4 mL de agua destilada, 1 mL de Coomassie 1 g/L preparado en HCl 2,2% v/v y 100  $\mu$ L de muestra o proteína estándar de calibración. Previamente se realizó una curva de calibración con BSA 0,5 mg/mL en concentraciones entre 0,016 y 0,5 mg/mL. Se midió la absorbancia a 465 nm ( $OD_{465}$ ) y 595 nm ( $OD_{595}$ ) y se calculó la razón  $O.D_{595}/O.D_{465}$ . El valor obtenido en ausencia de BSA se utilizó como blanco. Finalmente se graficó en función de la concentración de BSA.

## 4.4 RESULTADOS Y DISCUSIÓN

### 4.4.1 Mutación de la triada catalítica de la lipasa 2-17

#### 4.4.1.1 Clonamiento de los mutantes para la triada catalítica de la lipasa 2-17

En este trabajo, se clonaron y expresaron cinco mutantes del gen de la lipasa 2-17 activa a bajas temperaturas con el objetivo de identificar los aminoácidos de la triada catalítica. Para obtener cada uno de los mutantes los fragmentos anteriores y posteriores de cada una de las mutaciones fueron amplificados mediante PCR a partir del gen de la lipasa 2-17 clonado previamente en el vector de expresión pMALc -2E (Figura 50).



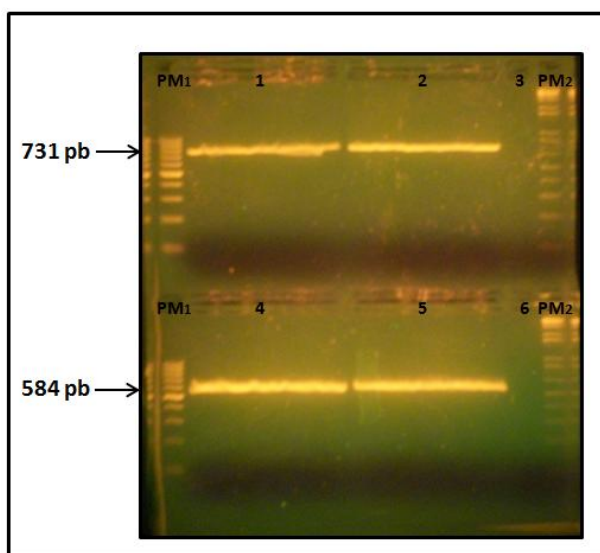
**Figura 50: Mutagénesis sitio dirigida.** Esquema general para el proceso mutagénico realizado en la lipasa 2-17. Las hebras nativas, las mutadas y los partidores A - B son ilustrados en celeste, púrpura y verde respectivamente. El partidor M, corresponde al partidor con la mutación y es ilustrado en púrpura, al igual que las hebras mutadas.

Para ello se diseñaron partidores específicos para cada una de las mutaciones, lo que permitió que tanto el fragmento directo como el inverso de cada mutación se complementaran para formar el fragmento completo final del gen con cada mutación con un tamaño esperado de aproximadamente 1300 pb. Para ello se utilizó la enzima KOD® ya que esta enzima repara errores en la amplificación dada su actividad *proofreading*. Las mutaciones realizadas correspondieron al cambio de la Serina 239 por una Asparagina, del Aspartato 361 por una Alanina, una Asparagina y un Glutamato y finalmente de la

Histidina 391 por una Asparagina, según las recomendaciones y los experimentos realizados por Jager S. (1992), Brumlik y Buckley, (1996).

- **AMPLIFICACIÓN FRAGMENTOS MUTACIÓN S239N**

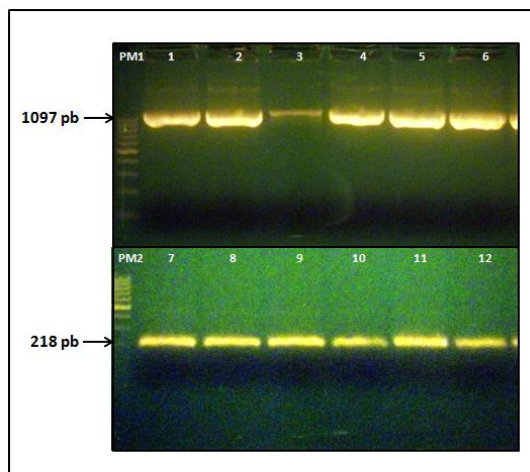
La mutación de la serina 239 fue realizada mediante dos PCR iniciales (Figura 51). El primero utilizó los partidores LipFmut y LipS239NR que amplificaron el fragmento directo desde el inicio del gen hasta la mutación, mientras que el segundo PCR utilizó los partidores LipS239NF y LipRmut que amplificaron el fragmento inverso desde el final del gen hasta la mutación. Los partidores fueron diseñados para dejar una superposición de 19 pares de bases. El control negativo de cada uno de los PCRs no presentó producto de amplificación (datos no mostrados).



**Figura 51: PCR 1 y 2 de la mutación S239N. PM1:** Marcador de peso molecular 100 bp. **PM2:** Marcador de peso molecular 1Kb. **Carril 1:** Fragmento directo S239N. **Carril 2:** Duplicado Fragmento directo S239N. **Carril 3:** Control negativo fragmento forward S239N. **Carril 4:** Fragmento inverso S239N. **Carril 5:** Duplicado Fragmento inverso S239N. **Carril 6:** Control negativo fragmento inverso S239N.

- **AMPLIFICACIÓN FRAGMENTOS MUTACIÓN D361A, D361E, D361N:**

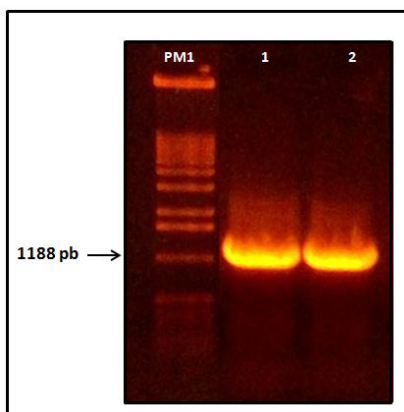
La mutación de Aspartato 361 se realizó de la misma manera que la descrita para la Serina 239. Se utilizaron los partidores LipFmut y LipRmut para los extremos y partidores con cada una de las mutaciones, alanina, glutamato y arginina para cada caso. La superposición de los fragmentos fue de 20 pares de bases, para aumentar la probabilidad de complementación correcta entre los fragmentos (Figura 52). El control negativo de cada uno de los PCRs no presentó producto de amplificación (datos no mostrados).



**Figura 52: PCR 1 y 2 de las mutaciones D361A, D361E y D361N.** Fragmento directo 1097 pb mutación D361A, D361E y D361N. **PM1:** marcador de peso molecular 1Kb. **Carril 1:** Fragmento directo D361A. **Carril 2:** Duplicado Fragmento directo D361A. **Carril 3:** Fragmento directo D361E. **Carril 4:** Duplicado Fragmento directo D361E. **Carril 5:** Fragmento directo D361N. **Carril 6:** Duplicado Fragmento directo D361N. Control negativo no presentó amplificado (Datos no mostrados). Fragmento inverso 218 pb mutación D361A, D361E y D361N. **PM2:** Peso Molecular 1Kb. **Carril 7:** Fragmento inverso D361A. **Carril 8:** Duplicado Fragmento inverso D361A. **Carril 9:** Fragmento inverso D361E. **Carril 10:** Duplicado Fragmento inverso D361E. **Carril 11:** Fragmento inverso D361N. **Carril 12:** Duplicado Fragmento inverso D361N. Control negativo no presentó amplificado (Datos no mostrados).

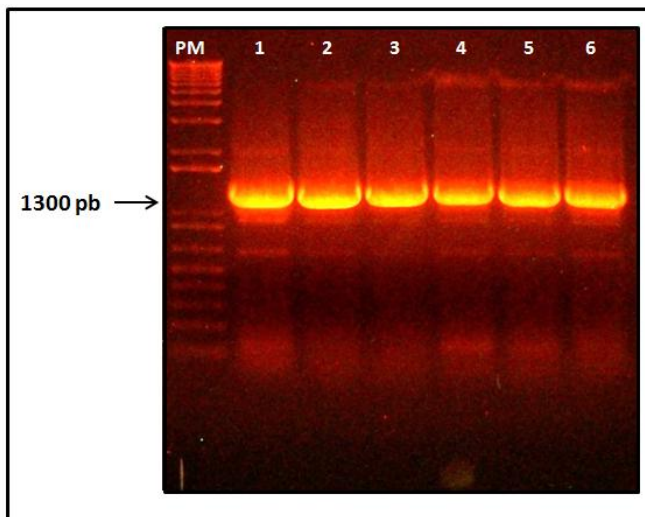
- **AMPLIFICACIÓN FRAGMENTOS MUTACIÓN H391N**

Finalmente, para la mutación de la Histidina catalítica, se utilizaron los partidores LipFmut y LipH391NR para obtener el fragmento directo y los partidores LipH391NF y LipRmut para obtener el inverso. Al igual que en las mutaciones anteriores la superposición de los fragmentos para el segundo PCR fue de 20 pares de bases (Figura 53). El control negativo de cada uno de los PCRs no presentó producto de amplificación (datos no mostrados).



**Figura 53: PCR 1 fragmento directo H391N.** **PM1:** marcador de peso molecular 1 Kb. **Carril 1:** Fragmento directo H391N. **Carril 2:** Duplicado fragmento directo H391N. Los controles negativos se realizaron y resultaron no exhibir bandas (datos no mostrados).

Una vez realizado cada uno de los PCR 1 y 2 para cada una de las mutaciones, se procedió a realizar un PCR para unir cada fragmento directo con su correspondiente fragmento inverso. Los productos de PCR resultantes para cada una de las mutaciones fueron de un tamaño aproximado de 1300 pb, los cuales fueron cortados desde un gel de agarosa 1 %, purificados y ligados en el vector pGEM-T y directamente al vector de expresión pMALc-2E (Figura 54).

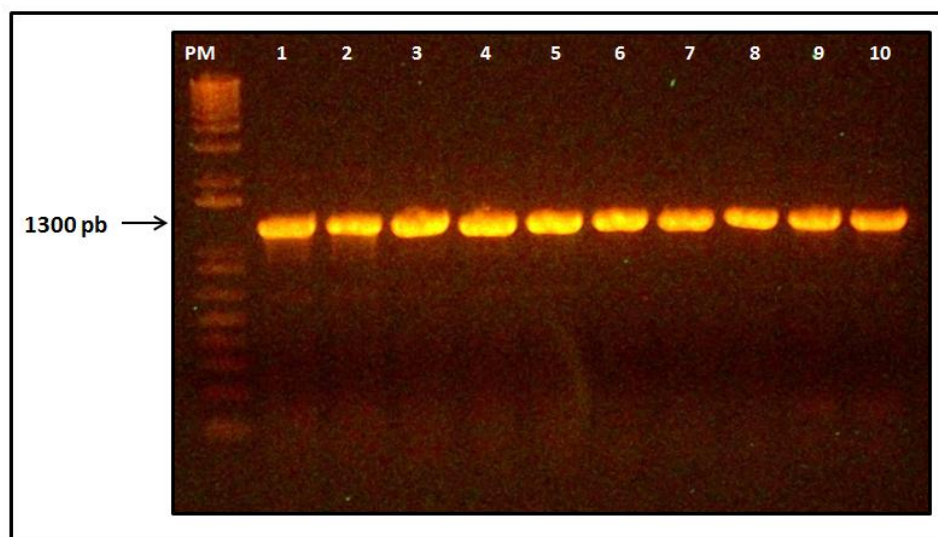


**Figura 54: PCR final para cada mutación. PM:** Marcador de peso molecular 1Kb. **Carril 1:** Fragmento gen completo lipasa 2-17 sin mutar. **Carril 2:** Fragmento gen completo D361A. **Carril 3:** Fragmento gen completo D361E. **Carril 4:** Fragmento gen completo D361N. **Carril 6:** Fragmento gen completo S239N. **Carril 8:** Fragmento gen completo H391N.

El producto de ligación fue utilizado para transformar células *E. coli* DH5 $\alpha$  electrocompetentes, que se crecieron en placas con medio LB-Agar suplementado con ampicilina 100  $\mu$ g/mL, y se incubaron a 37  $^{\circ}$ C por 16 h. Se comprobó la presencia de inserto en el vector, en cada una de las mutaciones, mediante PCR de las colonias obtenidas, utilizando los partidores LipFmut y LipRmut (datos no mostrados) y se extrajo el ADN plasmidial de las colonias que resultaron ser positivas.

Estas colonias se crecieron en medios TB y LB suplementado con ampicilina 100  $\mu$ g/mL y los productos de las minipreparaciones se utilizaron para transformar células *E. coli* BL21 DE3 electrocompetentes que posteriormente fueron crecidas en placas con medio LB-Agar suplementadas con ampicilina, X-Gal e IPTG, y se incubaron a 37  $^{\circ}$ C por 16 h. Nuevamente se realizaron PCR de colonias para corroborar la presencia del inserto en cada una de las mutaciones (Figura 55). Se seleccionaron los clones positivos en cada caso los cuales se enviaron a ser secuenciados (ver secuencia en Anexos) y se utilizaron para continuar con la expresión de los mutantes.



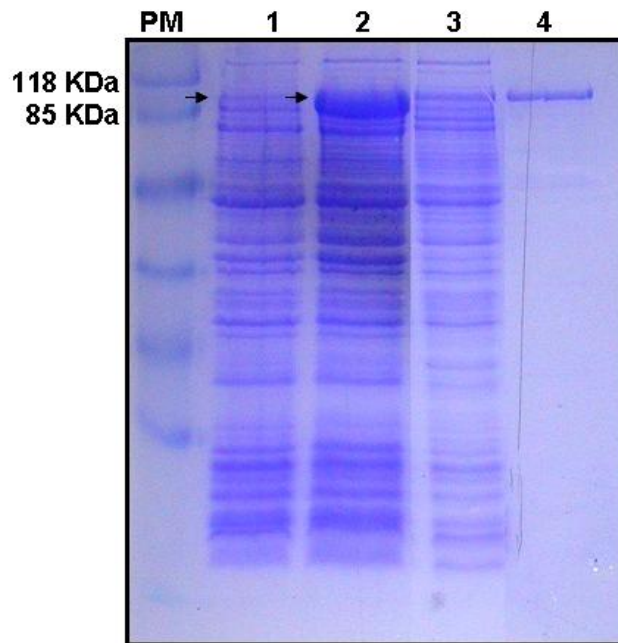


**Figura 55: Electroforesis de PCR de colonias.** Electroforesis en gel de agarosa al 1% teñido con bromuro de etidio. Carriles desde el 1 al 10 muestran PCR de colonias para corroborar la presencia del inserto en el vector pMAL-c2E de cada una de las mutaciones. **PM:** Marcador de peso molecular 1Kb. **Carril 1 y 2:** Fragmento D361A. **Carril 3 y 4:** Fragmento D361E. **Carril 5 y 6:** Fragmento D361N. **Carril 7 y 8:** Fragmento S239N. **Carril 9 y 10:** Fragmento H391N.

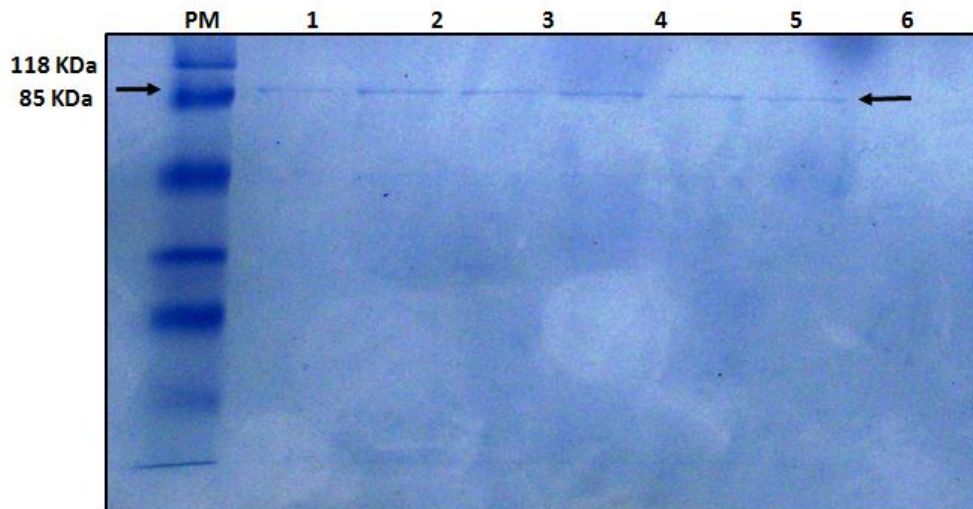
#### 4.4.1.2 Expresión de los mutantes para la triada catalítica de la lipasa 2-17

Luego de seleccionar los clones de *E. coli* BL12 DE3 transformados con el vector pMAL-c2E más el gen de la lipasa 2-17 mutado para cada uno de los aminoácidos de la triada, se indujo la expresión de los genes con IPTG 0,3 mM, como está descrito en la sección 4.3.2.10 de Materiales y Métodos. El sistema de expresión pMAL-c2E genera proteínas quiméricas que contienen un dominio de unión a maltosa (MBP). Esto facilita el proceso de purificación ya que permite la utilización de una columna de afinidad a amilosa y frecuentemente estabiliza la proteína recombinante (Baneyx, 1999). Por otra parte, una vez expresadas las proteínas se realizó lisis celular en cada caso, mediante sonicación para obtener en el extracto crudo la lipasa de interés, ya que el vector pMAL-c2E tiene una delección en la secuencia señal *malE*, lo que resulta en la expresión citoplasmática de la proteína de fusión.

Los extractos crudos de cada mutación se purificaron mediante una columna de amilosa y finalmente se recolectaron fracciones de 4 mL con la proteína eluída. La fracción citoplásmica y el extracto crudo obtenidos en cada caso, se analizaron en geles de poliacrilamida, donde se observaron altos niveles de expresión de la proteína de fusión. Las lipasas quiméricas presentaron un tamaño de 90 kDa, lo que es consistente con la suma de la lipasa de 47 kDa y de la MBP de 43 kDa (Figura 56). Se lograron obtener las proteínas de fusión puras para cada una de las mutaciones (Figura 57).



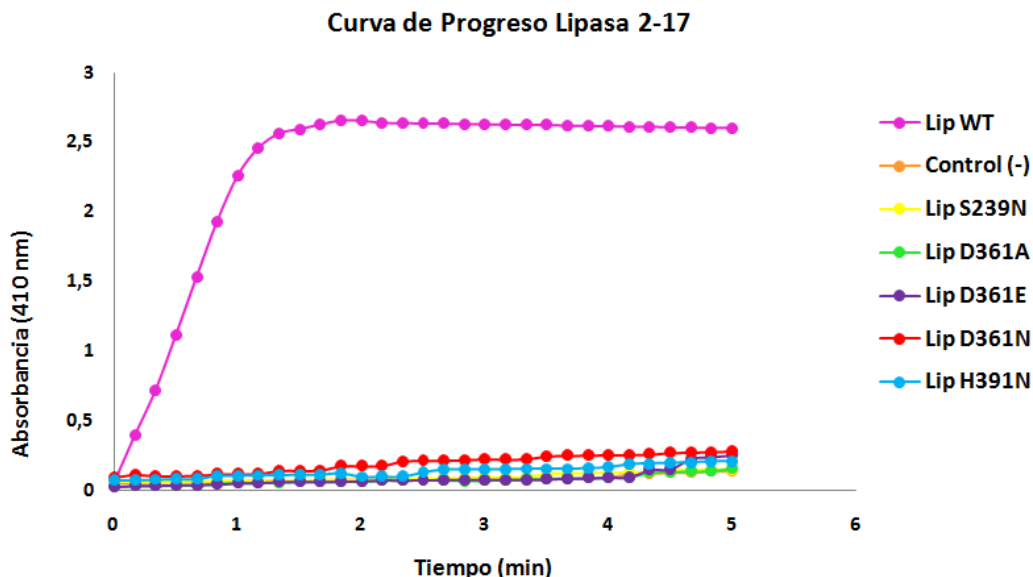
**Figura 56: Expresión del mutante S239N de la lipasa 2-17.** PM: Marcador de peso molecular. **Carril 1:** Extracto crudo de células sin inducir. **Carril 2:** Células inducidas con 0,3 mM IPTG, **Carril 3:** Sobrenadante fracción soluble una vez sonicadas, **Carril 4:** Purificación de la proteína con resina de amilosa. Las flechas indican la banda correspondiente a la lipasa quimérica.



**Figura 57: Lipasa quimérica purificada para cada una de las mutaciones.** PM: Marcador de peso molecular. **Carril 1:** Purificación de la lipasa nativa con resina de amilosa. **Carril 2:** Purificación de la lipasa con la mutación S239N. **Carril 3:** Purificación de la lipasa con la mutación D361A. **Carril 4:** Purificación de la lipasa con la mutación D361N. **Carril 5:** Purificación de la lipasa con la mutación D361E. **Carril 6:** Purificación de la lipasa con la mutación H391N.

#### 4.4.1.3 Parámetros cinéticos de las mutantes de la triada catalítica

Finalmente la cinética de reacción se estableció mediante una curva de progreso a 20°C, utilizando una concentración de sustrato de 0,1 M. El aumento de la absorbancia a 410 nm en el tiempo demostró que la enzima recombinante se encontraba activa mientras que los mutantes presentaron pérdida de actividad (Figura 58).



**Figura 58: Curva de progreso de la lipasa 2-17 y de los mutantes a temperatura a 20°C.** Los ensayos de actividad lipolítica se realizaron con el sustrato *p*-nitrofenil caprato como fue descrito en la sección 4.3.2.13, durante un tiempo de incubación de 5 minutos. Las líneas púrpura y anaranjada corresponden al ensayo realizado con y sin la enzima nativa, respectivamente. Las líneas amarilla, verde, morada, roja y celeste corresponden a los mutantes S239N, D361A, D361E, D361N y H391N respectivamente.

Mediante el método de Bradford se determinaron las concentraciones de proteína total obtenidas en cada uno de los casos (sección 4.3.2.14). Para determinar aproximadamente cuál era la concentración de lipasa una vez purificada la proteína de fusión, se estableció una relación porcentual entre los pesos moleculares de las dos proteínas, la lipasa 2-17 de 47 KDa y la MBP de 43 KDa. Debido a que el porcentaje de la lipasa es de un 52,22% con respecto a la proteína total, la concentración de lipasa obtenida en cada caso fue de 1,16 mg/mL para la lipasa nativa, de 1,04 mg/mL para la lipasa mutante S239N, de 0,95 mg/mL para la lipasa mutante D361A, de 1,18 mg/mL para la lipasa mutante D361E, de 1,05 mg/mL para la lipasa mutante D361N, de 1,02 mg/mL para la lipasa mutante H391N.

Se definió una unidad de actividad lipolítica (U) como  $\mu$ moles *p*-nitrofenol/min formados. Para ello, se realizó una curva de calibración para diferentes concentraciones del producto *p*-nitrofenol a 20°C y se midió la absorbancia a 410 nm (ver sección 4.3.2.13). En la curva de calibración se observó una pendiente de 0,153 unidades de absorbancia \* L/ mmol de *p*-nitrofenol en un volumen final de reacción

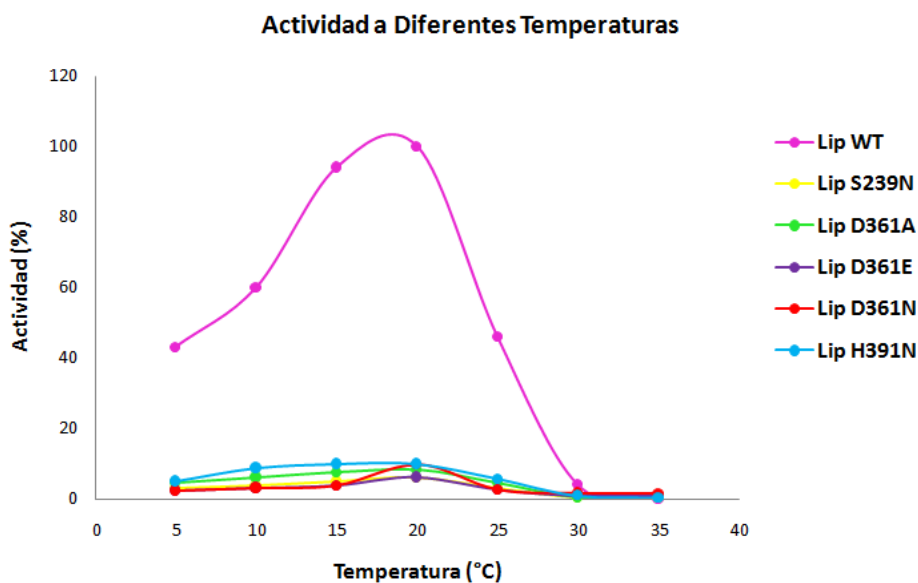
de 1 mL, y se utilizaron para dividir las unidades de absorbancia/min calculadas en la primera parte (lineal) de la curva de progreso de la lipasa 2-17. Se obtuvieron 15,88  $\mu$ moles/min de *p*-nitrofenol (U). Finalmente para tener una aproximación de la actividad específica de la lipasa en cada caso, se dividió este valor por la cantidad de enzima agregada en la reacción. Los resultados se muestran en la tabla VIII.

**Tabla VIII: Propiedades de cada una de las lipasas mutantes comparadas con la lipasa nativa.**

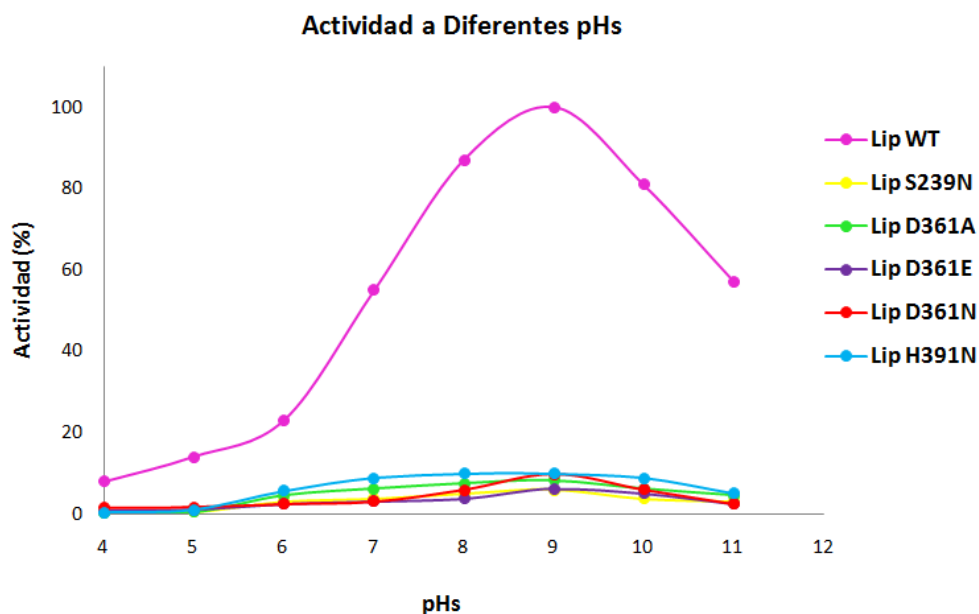
Mutación	Cambio de Codón	Concentración de Proteína [mg/mL]	Actividad Específica [ $\mu$ mol/min*mg]
Nativa	-	1,16	136,9
S239N	AGT a AAC	1,04	-
D361A	GAT a GCC	0,95	-
D361E	GAT a GAG	1,18	-
D361N	GAT a AAC	1,05	-
H391N	CAT a AAC	1,02	-

\* - no presentan actividad

Para descartar que los mutantes no hubiesen cambiado su dependencia de la temperatura (Figura 59) y pHs (Figura 60) a las cuales se obtiene su máxima actividad, se realizó un ensayo de cada uno de los mutantes a diferentes temperaturas y pHs.



**Figura 59: Determinación de actividad a diferentes temperaturas.** Se determinó la actividad de los mutantes a diferentes temperaturas con el fin de corroborar que la mutación no alterara la estabilidad térmica. Los ensayos de actividad lipolítica se realizaron con el sustrato *p*-nitrofenil caprato como fue descrito en la sección 4.3.2.13. La línea púrpura corresponde al ensayo realizado con la enzima nativa. Las líneas amarilla, verde, morada, roja y celeste corresponden a los mutantes S239N, D361A, D361E, D361N y H391N respectivamente. El control negativo no presentó hidrólisis del sustrato (datos no mostrados).

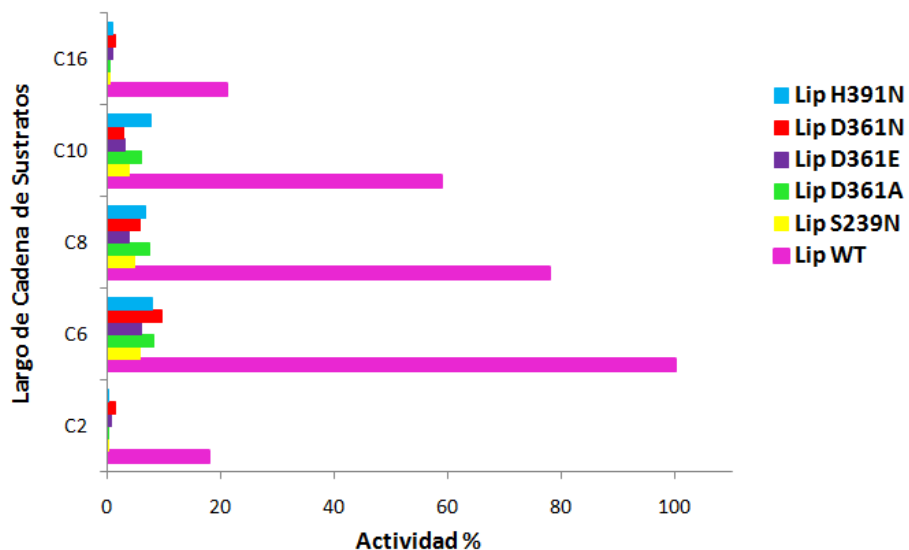


**Figura 60: Determinación de actividad a diferentes pHs.** Se determinó la actividad de los mutantes a diferentes pHs con el fin de corroborar que el pH al cual se obtiene el máximo de actividad no hubiese cambiado. Los ensayos de actividad lipolítica se realizaron con el sustrato p-nitrofenil caprato como fue descrito en la sección 4.3.2.13. La línea púrpura corresponde al ensayo realizado con la enzima nativa. Las líneas amarilla, verde, morada, roja y celeste corresponden a los mutantes S239N, D361A, D361E, D361N y H391N respectivamente. El control negativo no presentó hidrólisis del sustrato (datos no mostrados).

Finalmente se evaluó el efecto de la mutación en la especificidad de sustrato de los variantes, debido principalmente a las modificaciones espaciales del sitio activo. Las mutaciones, además de alterar las propiedades de la enzima, pueden rara vez influir en la selectividad por un determinado sustrato. Para descartar que las mutantes no cambiaran esta selectividad, se realizaron las curvas de progreso con sustratos de cadenas de 2, 6, 8, 10 y 16 carbonos, utilizando como control la lipasa nativa (Figura 61).

Según los resultados obtenidos, ninguno de los variantes presentó actividad en el rango de temperatura y pH, donde se obtuvo el máximo de actividad para la lipasa 2-17. Debido a que todos los variantes se produjeron en cantidades similares, es posible atribuir ésta pérdida de actividad a la variación en la triada catalítica y no a la falta de proteína en la muestra. Además al analizar la especificidad de sustrato, fue posible ver que ninguno de los mutantes cambió su selectividad por un sustrato de menor o de mayor largo de cadena. Esto podría indicar la importancia que tiene cada uno de éstos aminoácidos en la catálisis, y respaldar la hipótesis de que la posible triada catalítica de la lipasa 2-17 corresponda a la Serina 239, Aspartato 361 e Histidina 391.

### Actividad sobre Diferentes Sustratos



**Figura 61: Determinación de actividad catalítica de la lipasa sobre diferentes sustratos.** Se midió la actividad de los variantes sobre diferentes sustratos con el fin de corroborar que la mutación no alterara la especificidad de sustrato. Los ensayos de actividad lipolítica se realizaron con el sustrato p-nitrofenil caprato como fue descrito en la sección 4.3.2.13. La barra púrpura corresponde al ensayo realizado con la enzima nativa. Las barras amarilla, verde, morada, roja y celeste corresponden a los mutantes S239N, D361A, D361E, D361N y H391N respectivamente.

## 4.5 CONCLUSIONES

Si bien son las cadenas laterales las que definen las características fisicoquímicas de los aminoácidos (como polaridad, hidrofobicidad, aromaticidad, flexibilidad conformacional, tamaño, etc), existen equivalencias a la hora de influir en el plegamiento de proteínas, ya que en gran medida no dependen estrictamente de cada uno de los aminoácidos, sino mas bien dependen del equilibrio entre las fuerzas existentes al momento de realizarse el plegamiento. Así, es posible encontrar aminoácidos que son estructuralmente intercambiables como es el caso del grupo de aminoácidos Asp, Glu, Asn y Gln o bien Ser, Thr, Asn y Gln. Sin embargo, a la hora de influir en la actividad de la proteína la realidad es otra. Así, para cada una de las mutaciones evaluadas en este capítulo, los resultados obtenidos no fueron satisfactorios en términos de la hidrólisis del sustrato.

La determinación de la primera estructura tridimensional de una quimiotripsina permitió tener una visión más profunda del mecanismo de acción de un gran número de enzimas. Esto demostró que la serina y la histidina están en una posición adecuada para funcionar de acuerdo al mecanismo catalítico tipo triada. El oxígeno de la serina y el nitrógeno de la histidina, se encuentra a una distancia tal, que les permite formar un puente de hidrógeno. El otro nitrógeno del anillo imidazol está unido al hidrógeno del aspartato, de modo que su enlace de hidrógeno se encuentra protegido del agua por varios residuos aminoácidos.

La relación geométrica entre el aspartato, la histidina y la serina ha llevado a la hipótesis de que la histidina sirve para transferir un protón desde la serina al aspartato en un mecanismo de relevo de cargas. Sin embargo este “relevo de protones” desde el grupo OH altamente básico de la serina hasta el aspartato ácido es químicamente poco probable. Otra hipótesis apunta a que el aspartato puede estar involucrado en la estabilización del par iónico generado entre el ion de grupo imidazol y el intermediario tetraédrico cargado negativamente, y que el aspartato puede participar en la orientación correcta del tautómero de la histidina en relación a la serina.

La interacción entre el aspartato y el grupo imidazol es más importante en el estado de transición de la formación del intermediario tetraédrico que en el estado fundamental de la reacción. Estudios (Jordan & Polgar, 1981), demostraron que mientras que el protón aun esta en el grupo imidazol, el enlace de hidrógeno con el aspartato es mucho más fuerte en el estado de transición que en la enzima nativa.

Otros estudios han reportado que la sustitución del ácido aspártico catalítico por una asparagina resulta en una enzima mutante cuya relación  $k_{cat}/K_m$  es sobre  $10^4$  veces menor que la de la enzima nativa (Craik y col., 1987), lo que fue concordante con los resultados obtenidos en el capítulo anterior. Por otro lado, la determinación de la estructura tridimensional de los variantes indicó que la histidina catalítica fue incapaz de aceptar el protón de la serina. El análisis cinético realizado por Craik y col.,

(1987) también demostró que la mutación reducía en gran medida la  $k_{cat}$ , y que alteró levemente la  $K_m$ . Específicamente uno de los nitrógenos de la histidina (NE2) fue protonado, mientras que el no protonado formó un enlace de hidrógeno con el grupo amida de la asparagina. Cuando ambos nitrógenos están protonados, la histidina adopta una conformación alternativa e interactúa con una molécula de agua en lugar de interactuar con el residuo de asparagina.

El rol de la serina y de la histidina en la catálisis es bastante claro. La serina forma el intermediario acil-enzima y la histidina transfiere el protón desde el grupo OH de la serina al grupo saliente del sustrato para formar el complejo acil-enzima. De una forma similar, la histidina también opera en la hidrólisis del complejo. Esto nuevamente fue concordante con los resultados obtenidos previamente en los cuales, se realizó el cambio de la serina y la histidina (Corey & Craik, 1992), y cuyas mutantes exhibieron una disminución de los valores de  $k_{cat}/K_m$   $10^5$  veces en comparación a las enzimas nativas. Además se demostró que mutaciones particulares de serina o histidina anulaban la actividad de la enzima de manera similar que en el caso de las mutantes dobles, indicando que solo una mutación suprime la reacción hidrolítica normal.



## **CAPÍTULO 5**

**Estudio del Fenómeno de Activación Interfacial de la Lipasa 2-17**

## 5.1 RESUMEN

Finalmente en este capítulo, para conocer el comportamiento y las interacciones de la triada, el sustrato y el resto de la proteína, se comparó la estructura de la lipasa 2-17 con lipasas cuya activación interfacial está mediada por estructuras que actúan como tapa. Particularmente se realizó la comparación con la lipasa M37 (2ORY) cuya hélice  $\alpha_3$ , compuesta por los aminoácidos Ser96, Ile97, Ser98, Asp 99, Trip100 y leu101, bloquea el sitio activo dejándolo inaccesible al sustrato. Esta hélice que actúa como tapa tiene una composición anfifílica, con la zona hidrofóbica hacia el interior de la enzima y la hidrofílica hacia el exterior.

Para determinar cuáles son las posibles estructuras secundarias que podrían estar actuando como tapa en la lipasa 2-17, se realizó un alineamiento estructural entre la lipasa 2-17 y la lipasa M37. Este alineamiento mostro que tres estructuras podrían ser las encargadas de mediar el fenómeno de activación interfacial. Se definieron entonces, las estructuras  $\alpha_1$ ,  $\alpha_8$  y G2, como las posibles tapas de la lipasa 2-17. Estructuralmente estas hélices se encuentran sobre el sitio activo, pero en una conformación abierta ya que la estructura molde definida para realizar el modelamiento, se encontraba abierta.

Para conocer las propiedades anfifílicas de las hélices, se analizaron las composiciones de cada una de estas estructuras, y se observó que solo dos de ellas,  $\alpha_8$  y G2, presentaban esta codición a lo largo de su estructura.

Así, fue posible establecer que en la lipasa 2-17 dos estructuras tipo hélices  $\alpha$ ,  $\alpha_8$  y G2 que se encuentran ubicadas en el acceso al sitio activo, podrían ser las estructuras tipo tapa de esta enzima, lo que podría indicar que la lipasa 2-17 presenta el fenómeno de activación interfacial.

## 5.2 INTRODUCCIÓN

Sarda y Desnuelle en 1958 observaron que la lipasa que estaban estudiando, a diferencia de otras esterasas, era particularmente activa cuando el sustrato se encontraba formando una emulsión, pero esta actividad disminuía notablemente cuando el sustrato se encontraba disuelto. Este fenómeno entonces fue denominado activación interfacial y hoy en día ha permitido caracterizar y distinguir a las lipasas verdaderas de las esterasas.

Desnuelle y col., (1960) propusieron que para explicar este fenómeno, estaría implicado un cambio conformacional en la enzima. Esta teoría fue validada con la determinación de la primera estructura tridimensional de este tipo de enzimas, la cual exhibió la presencia de una estructura tipo  $\alpha$  hélice anfifílica ubicada en la superficie de la proteína, que cubría el sitio activo cuando la enzima se encontraba en medio acuoso, impidiendo la entrada del sustrato dentro de la lipasa y por lo tanto, evitando su hidrólisis. Se propuso además que existe un reajuste de esta estructura cuando la enzima se encuentra en un medio lipídico permitiendo que el sitio activo quede accesible para la entrada del sustrato.

Stobircka y col., en 1998 evidenciaron la variación de la fluorescencia de un triptófano ubicado en la estructura tipo tapa de una lipasa, que era producida por el cambio entre las conformaciones abierta y cerrada de la enzima. Tres años después, Neves-Petersen y col., (2001) estudiaron la distribución de aminoácidos polares y apolares en la superficie de diversas lipasas en función del pH, observando que la región cercana al sitio activo era más hidrofóbica cuando estas enzimas se encontraban en una conformación abierta.

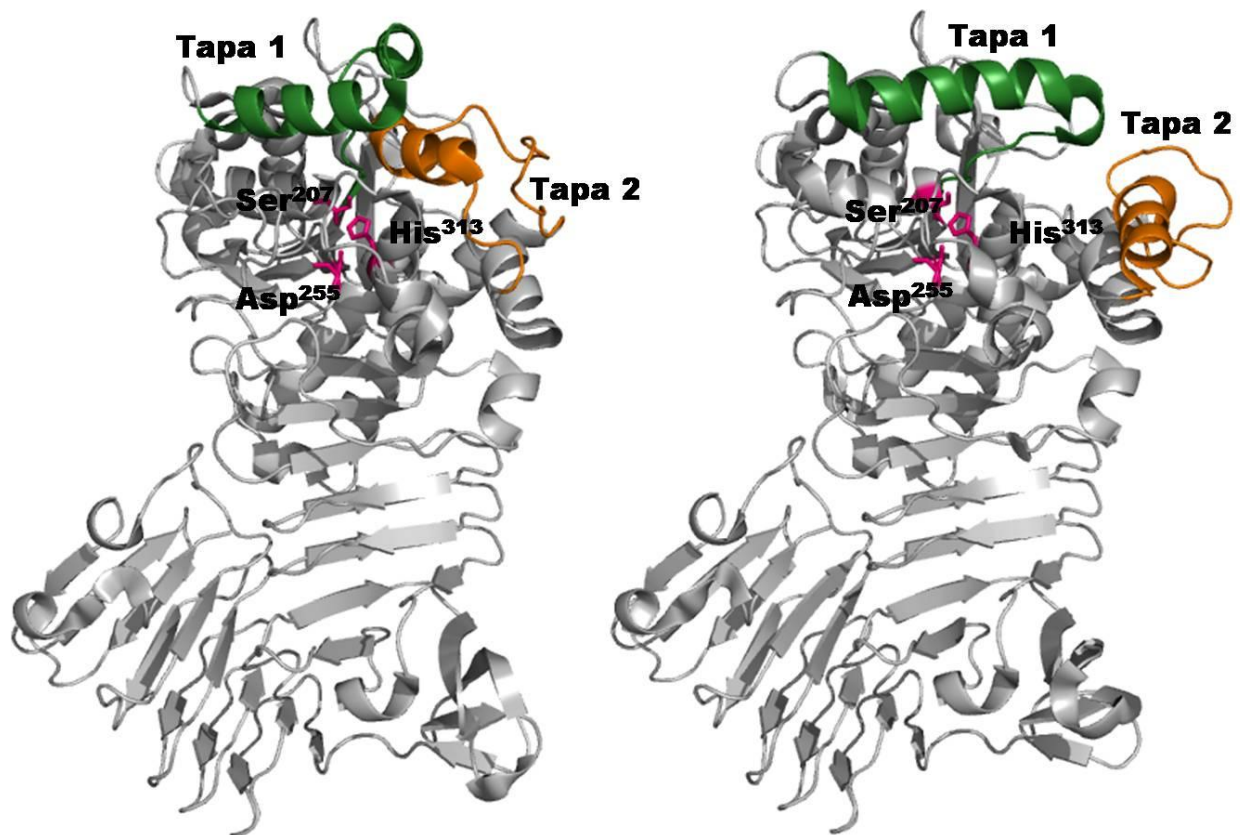
### 5.2.1 Activación interfacial

Las enzimas lipolíticas se caracterizan por su drástico incremento en la actividad cuando se encuentran en una interface lípido-acuosa. Este incremento en la actividad enzimática es provocado por rearrreglos estructurales en la región del sitio activo de la lipasa. En ausencia de una interface lípido-acuosa, el sitio activo es cubierto por una estructura denominada “tapa”. Sin embargo, en la presencia de sustratos hidrofóbicos, esta tapa se abre, haciendo a los residuos catalíticos accesibles al sustrato y exponiéndolos a la superficie hidrofóbica del sitio activo. Esta superficie hidrofóbica podría estar interactuando con la interface lipídica. Generalmente, esta tapa está formada por una o dos  $\alpha$  hélices, o bien por una región compuesta por *loops* (Panaiotov & Verger, 2000)

Brzozowski y col., (1990) mediante la estructura cristalográfica de una lipasa extracelular obtenida de *Rhizomucor miehei*, determinaron que el movimiento de 15 aminoácidos a lo largo de la tapa

(residuos 82-96) es un movimiento rígido de tipo bisagra, que trasporta algunos de los átomos de una corta  $\alpha$  hélice (residuos 85-91) por más de 12Å. Además determinaron que existen dos regiones (residuos 83-84 y 91-95) en los cuales pronunciadas transiciones de estructura secundaria entre las conformaciones  $\alpha$  y  $\beta$  son causados por dramáticos cambios en la conformación específica de los ángulos dihedrales ( $\phi$  y  $\psi$ ).

Como resultado de este cambio, una zona hidrofóbica de 800 Å<sup>2</sup> (8% de la superficie total de la molécula) se expone. Otras lipasas también son conocidas por tener "tapas" similares a la observada en la enzima de *R. miehei*, lo que ha hecho posible estudiar la estereoquímica general de la activación de estas lipasas en interfaces lípido-acuosa mediante estudios estructurales de estas enzimas.



**Figura 62: Estructura cristalográfica de la lipasa PML. Izquierda:** La estructura cristalográfica de la lipasa PML en su conformación cerrada (PDB ID 2Z8X). **Derecha:** Lipasa PML en su conformación abierta (PDB ID 2ZVD). La lipasa se encuentra coloreada en gris, mientras que la tapa 1 para ambos casos esta coloreada en anaranjado y la tapa 2 en verde. Para las dos conformaciones, la triada catalítica se encuentra coloreada en rosa (Ser207, Asp 255 e His313).

Chen y col., (2012), realizaron una comparación de las estructuras de la lipasa PML en conformación abierta (2ZVD) y cerrada (2Z8X) indicando que esta enzima cuenta con dos tapas (tapas 1 y 2). Estas tapas cubren el sitio activo en la conformación cerrada y es movida lejos del sitio activo en la conformación abierta. La tapa 1 (Figura 62) se encuentra fuertemente doblada en la mitad de la  $\alpha$  hélice en la conformación cerrada, mientras que en la conformación abierta es una sola hélice organizada en línea recta, con sus residuos no polares hacia afuera y sus residuos polares hacia el interior.

### **5.3 MATERIALES Y MÉTODOS**

#### **5.3.1 Análisis estructural de la lipasa 2-17 para determinar la activación interfacial**

Las comparaciones, los alineamientos y los análisis de los modelos fueron hechas utilizando los programas Pymol, SPDB Viewer y Chimera. Cada uno de los estudios fue realizado con la proteína molde 1JJI y comparado con la proteína M37 (2ORY).

### **5.4 RESULTADOS Y DISCUSIÓN**

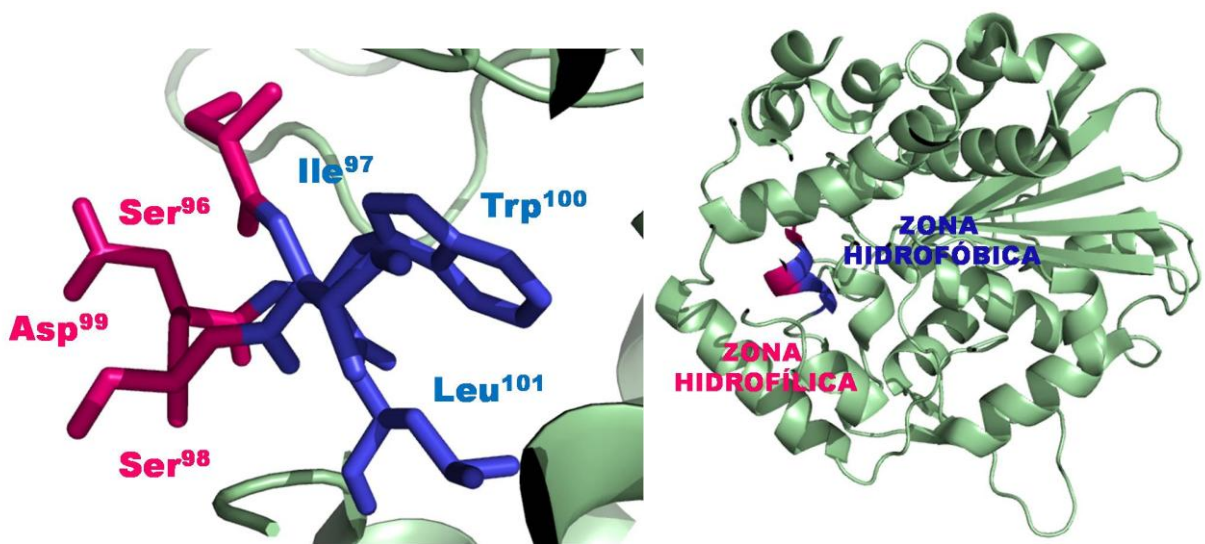
#### **5.4.1 Estudio de la activación interfacial de la Lipasa 2-17**

Se sabe que el mecanismo catalítico de la mayoría de las lipasas se basa en la reorganización de su estructura y que en presencia de interfaces hidrofóbicas, este dominio estructural conocido como “tapa” se somete a un cambio conformacional que expone la cavidad catalítica, permitiendo la comunicación con el sustrato y por lo tanto el comienzo de la actividad de la lipasa. El equilibrio entre conformaciones abiertas y cerradas es dependiente de las condiciones del medio, permitiendo que la tapa controle el acceso del sustrato al sitio activo.

Para determinar si la lipasa 2-17 presenta el fenómeno de activación interfacial, lo primero fue realizar una curva de progreso para evaluar la rapidez en la cual la enzima comienza a hidrolizar el sustrato. Según varios investigadores, la activación interfacial se ve reflejada en un pequeño retardo que las enzimas lipolíticas exhiben al comenzar la catálisis (Schrag y col., 1996; Cheng y col., 2012), y tal como es de esperarse, en la curva de progreso de la lipasa 2-17 es posible ver una pequeña demora de aproximadamente 35 segundos, antes de que comience a hidrolizarse el sustrato (ver figura en anexos 8.5.1).

Esto podría dar un pequeño indicio que la lipasa 2-17 podría tener estructuras secundarias que actúen como tapa, y que existe una reorganización cuando la enzima se encuentra en una interface lipídica. Posteriormente, se estudió la estructura tridimensional de la lipasa 2-17 y se comparó con la estructura tridimensional de una lipasa que tuviese estructuras secundarias que actúen como tapa y cuya activación interfacial es conocida (Figura 64).

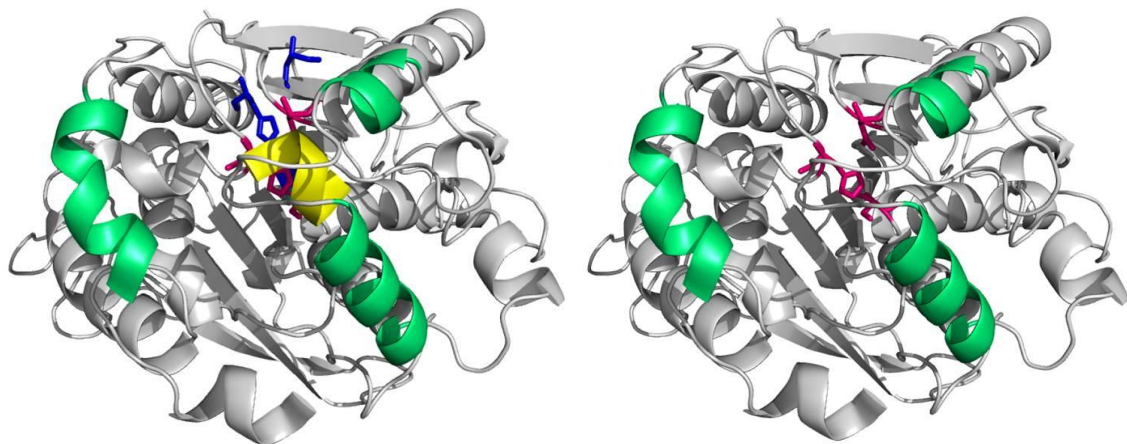
Particularmente, la comparación fue realizada con la lipasa M37 y se determinaron las estructuras secundarias que podrían estar participando en la formación de la tapa.



**Figura 63: Representación tapa  $\alpha 3$  de la lipasa psicrófila M37.** Lipasa M37 coloreada en verde. Aminoácidos hidrofóbicos (S96, S98, D99) coloreados en azul y aminoácidos hidrofílicos (I97, W100, L101) en rosa. Figura elaborada con el programa Pymol.

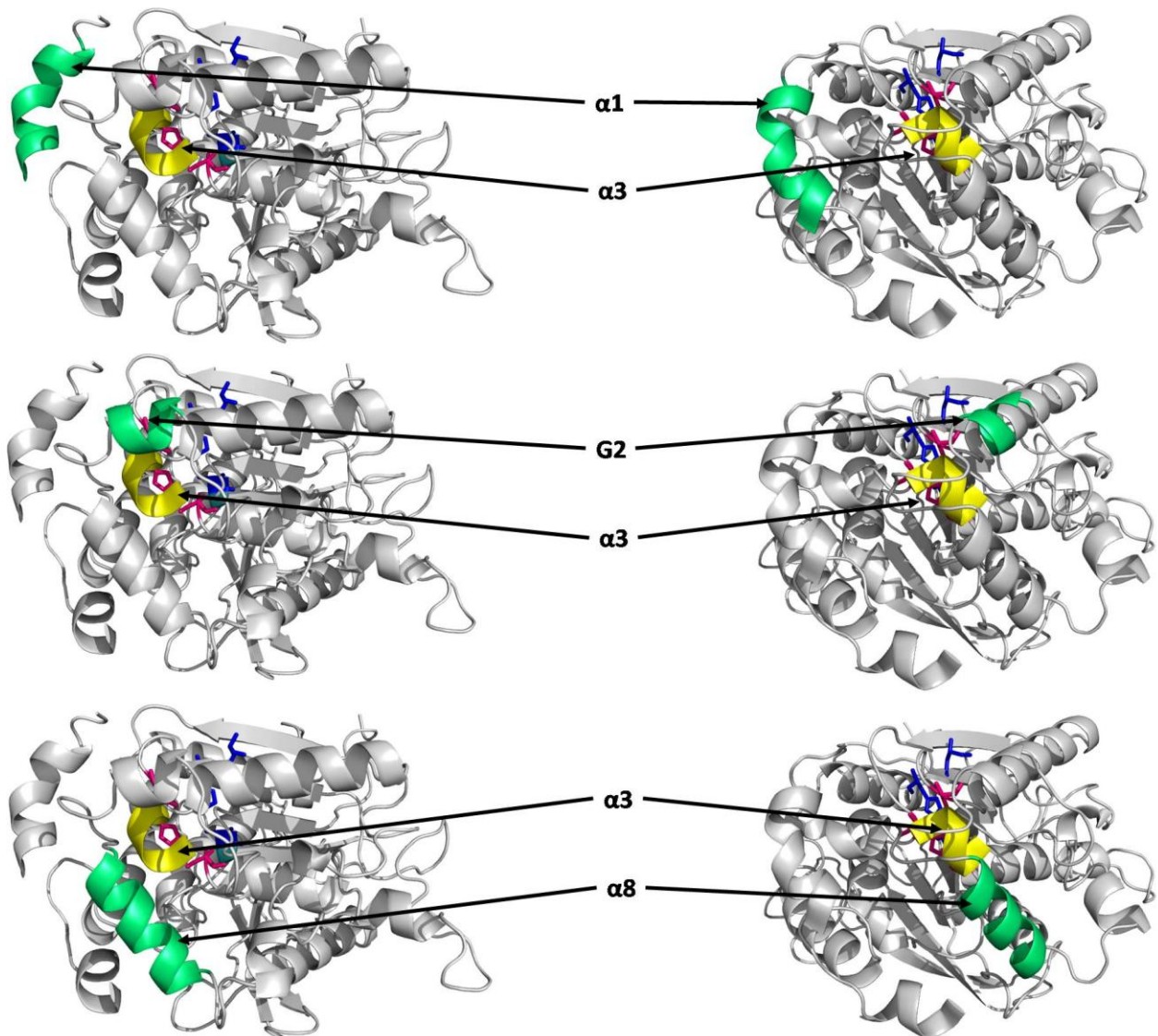
Jung y col., (2008) establecieron que en la lipasa M37 en solución acuosa, el sitio activo está completamente inaccesible al sustrato y que se encuentra cubierto por una  $\alpha$  hélice ( $\alpha 3$ ) que actúa como tapa. Esta hélice interactúa con la hoja  $\beta 2$  y su movimiento se ve influenciada por las  $\alpha$  hélices 9 y 10.

Se realizó un alineamiento estructural entre la lipasa 2-17 y la lipasa M37 (Figura 64), ya que a pesar de que sus estructuras presentan un porcentaje de identidad bastante bajo, el sitio activo espacialmente es bastante parecido.



**Figura 64: Comparación de estructuras tridimensionales que participarían como tapas en Lip 2-17 y M37 (PDB ID 2ORY).** Izquierda: Alineamiento estructural de la lipasa 2-17 con la lipasa M37. La estructura de la lipasa 2-17 se encuentra en color gris. Las triadas catalíticas de la lipasa 2-17 (rojo) y de la lipasa M37 (azul) fueron alineadas. La tapa de la lipasa M37 se encuentra coloreada de amarillo y las posibles tapas de la lipasa 2-17 en verde. Derecha: Estructura de la Lipasa 2-17. Figura elaborada con el programa Pymol.

El alineamiento reveló que existen tres posibles estructuras en la lipasa 2-17 que podrían estar actuando como tapa ya sea de manera independiente, o bien relacionadas entre sí. Estas estructuras corresponden a las hélices  $\alpha 1$  (L9, N10, A11, L12, L13, T14, K15, T16, L17, K18, T19, V20),  $\alpha 8$  (H317, N318, D319, A320, N321, V322, F323, N324, T325, A326, Y327) y G2 (P304, S305, W306, E307, L308) que se encuentran cubriendo la entrada al sitio activo de la enzima (Figura 65).

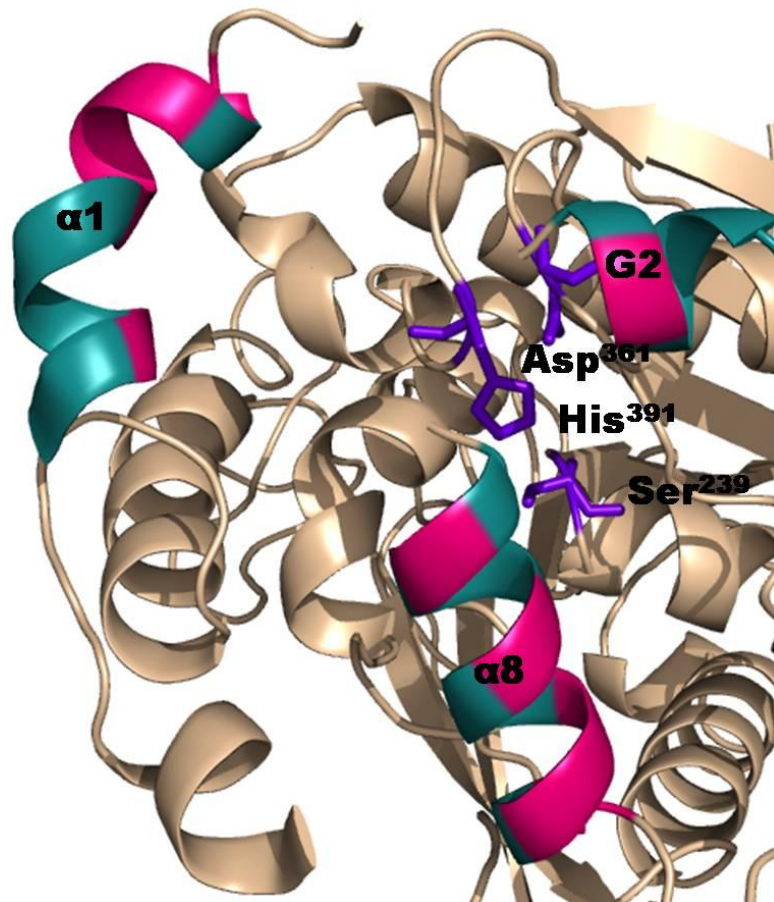


**Figura 65: Comparación de estructuras  $\alpha$  hélices de la lipasa 2-17 vistas de manera lateral (izquierda) y frontal (derecha), que podrían estar actuando como tapa.** Para todos los casos la lipasa 2-17 se encuentra coloreada con sus  $\alpha$  hélices y sus hojas  $\beta$  de color gris. La lipasa M37 fue alineada estructuralmente con la lipasa 2-17 y la hélice  $\alpha 3$  que actúa como tapa en esta enzima, se encuentra marcada con color amarillo. Las  $\alpha$  hélices que se encuentran coloreadas con verde, corresponden a las hélices  $\alpha 1$ , G2 y  $\alpha 8$ , candidatas como estructuras secundarias que podrían actuar como tapa en la lipasa 2-17. Figura elaborada con el programa Pymol.



Particularmente,  $\alpha 1$  se encuentra ubicada en un *loop* bastante flexible, y no interactúa directamente con otras estructuras, lo que podría favorecer el movimiento en presencia de una interface lipídica. Por otra parte, las estructuras  $\alpha 8$  y G2, se encuentran estratégicamente ubicadas por sobre la cavidad del sitio activo, lo que podría impedir la entrada del sustrato de manera eficiente en caso de presentarse el fenómeno de activación interfacial en esta proteína.

Según Cheng y col., (2012) las estructuras que actúan como tapa corresponden a hélices anfifílicas que controlan el acceso del sustrato al sitio activo, mediante el cambio conformacional al estar en contacto con el sustrato micelar. Para ello, las estructuras tipo tapa deben poseer una parte hidrofóbica que apunta hacia el interior de la proteína cuando la enzima se encuentra en estado cerrado, y una hidrofílica que apunta hacia afuera. Por lo tanto, se estudiaron las composiciones aminoacídicas de la tapa de la lipasa M37 y se compararon las hélices candidatas  $\alpha 1$ ,  $\alpha 8$  y G2 en la lipasa 2-17 (Figura 66).

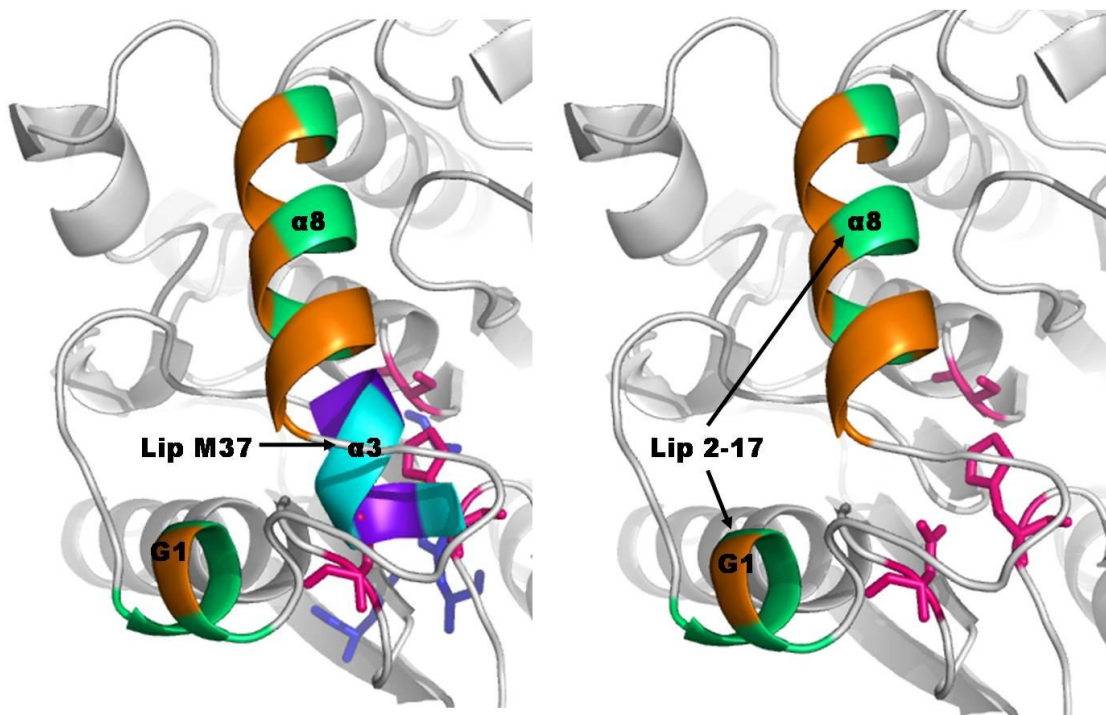


**Figura 66:** Modelo de la lipasa 2-17 con  $\alpha 1$ ,  $\alpha 8$  y G2. Estructura tridimensional de la lipasa 2-17 en beige. Estructuras  $\alpha$  hélices con  $\alpha 1$ ,  $\alpha 8$  y G2 coloreadas según afinidad por el agua. En rosa zonas hidrofílicas y en verde zonas hidrofóbicas. Triada catalítica indicada en morado. Figura elaborada con el programa Pymol.

La  $\alpha 3$  de la lipasa M37 contiene un 50% de aminoácidos apolares y un 50% de aminoácidos polares, ubicados en forma intercalada, lo que permite que a lo largo de toda la estructura  $\alpha$  hélice, una cara sea hidrofóbica y la otra hidrofílica.

Para el caso de la lipasa 2-17, la estructura  $\alpha 1$  dispone de 6 aminoácidos apolares y 5 polares, la  $\alpha 8$  de 4 aminoácidos apolares y 5 polares y G2 de 4 aminoácidos apolares y tan solo un aminoácido polar. Si bien podría pensarse que  $\alpha 1$  y  $\alpha 8$  tienen la composición equilibrada de aminoácidos hidrofílicos e hidrofóbicos y que G2 al presentar mayoritariamente aminoácidos hidrofóbicos no es un buen candidato, la distribución espacial de los aminoácidos en estas estructuras reveló que solo las estructuras  $\alpha 8$  y G2 presentaban una zona hidrofílica y una hidrofóbica a lo largo de la hélice, que podría favorecer el reordenamiento de la estructura para permitir la entrada del sustrato.

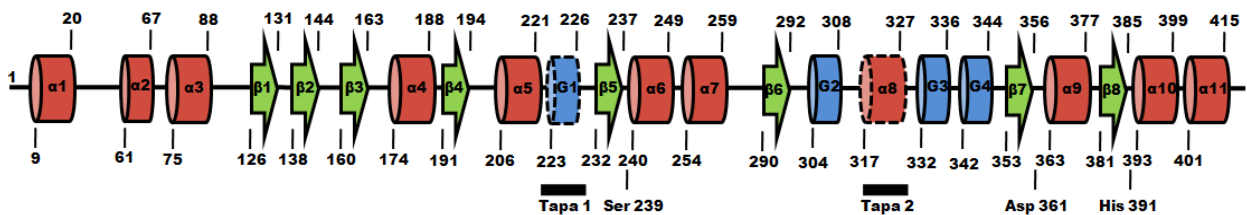
Finalmente se hizo una comparación directa entre la  $\alpha$  hélice de la lipasa M37 ( $\alpha 3$ ) y las dos candidatas de la lipasa 2-17 ( $\alpha 8$  y G2). Como es posible ver en la Figura 67, las estructuras se encuentran posicionadas de manera similar, con una orientación de sus aminoácidos hidrofílicos hacia el exterior de la molécula y sus aminoácidos hidrofóbicos hacia el interior, directamente en contacto con la cavidad del sitio activo (Figura 67).



**Figura 67: Esquema de la orientación de las estructuras que podrían actuar como tapa en la lipasa 2-17. Izquierda:** Superposición de las lipasas 2-17 y M37. La estructura  $\alpha 3$  se encuentra representada con sus aminoácidos hidrofílicos de color celeste y los hidrofóbicos de color morado. La triada catalítica de la lipasa M37 se encuentra de color azul. **Derecha:** Modelo de la lipasa 2-17. Para ambos casos en las estructuras  $\alpha 8$  y G2 los aminoácidos hidrofílicos e hidrofóbicos se encuentran representados en color anaranjado y verde respectivamente. La triada catalítica se representa en color rosa. Figura elaborada con el programa Pymol.

En el caso de que estas estructuras fueran las encargadas de permitir el acceso del sustrato al sitio activo, podría además existir una interacción entre  $\alpha 8$  y G2 dada su proximidad y su composición aminoacídica. Esta interacción entre ambas estructura podría permitir el movimiento de una o de las dos hélices, pero dado el tamaño de la cavidad, el movimiento de la hélice 8 podría ser suficiente para permitir la entrada del sustrato.

A pesar de que ambas estructuras se encuentran alejadas en la secuencia (Figura 68), y que una de ellas está compuesta por tan solo cinco aminoácidos, es común encontrar estructuras secundarias con estas características que actúan como tapa (Lopez-Gallegos y col., 2012).



**Figura 68: Representación esquemática de la estructura primaria de la lipasa 2-17.** Las  $\alpha$  hélices y las hojas  $\beta$  son representadas por cilindros y flechas respectivamente. Cada número indica la posición de cada residuo en relación a la metionina inicial en la proteína. Tanto las posibles tapas (1 y 2) como los aminoácidos de la triada catalítica son indicados en la figura.

Finalmente, al comparar las tapas seleccionadas para la lipasa 2-17 con las de otras lipasas, es posible ver que si bien existen discrepancias en los elementos estructurales que se encuentran alrededor, básicamente sus composiciones y ordenamientos aminoacídicos son muy similares. Al analizar los aminoácidos de la cavidad, es posible apreciar que aminoácidos hidrofóbicos como fenilalaninas, leucinas e isoleucinas se encuentran sobre el sitio activo sugiriendo que la unión al sustrato puede ser causada por estos residuos que son expuestos formando una amplia superficie hidrofóbica, proporcionando un ambiente apto para la catálisis.

## 5.5 CONCLUSIONES

Las lipasas en sistemas homogéneos se encuentran en una conformación cerrada, con el centro activo bloqueado, y en presencia de una interface hidrofóbica se adsorben a la misma, desplazando la estructura tipo tapa y originando la conformación abierta. Este mecanismo de acción permite a las lipasas actuar en interfaces. Dado que los sustratos naturales de las lipasas son grasas y aceites, esta actividad en la interface es un requerimiento indispensable para la función biológica de estas enzimas.

Sin embargo el hecho de que las lipasas en medio acuoso presenten actividad catalítica, ha permitido que algunos investigadores formulen la hipótesis que establece que las lipasas en sistemas acuosos homogéneos se encuentran en cierto equilibrio entre una conformación cerrada y una conformación abierta que permite que las lipasas sean catalíticamente activas en ausencia de interfaces.

Para el caso de la lipasa 2-17, se determinaron dos estructuras tipo hélices  $\alpha$ ,  $\alpha 8$  y G2, que podrían estar bloqueando el sitio activo de la enzima. Particularmente, la comparación se hizo con una estructura cuya conformación fue cerrada, para determinar la entrada al sitio activo, pero al encontrarse la lipasa 2-17 en una conformación abierta, fue necesario comparar las estructuras secundarias  $\alpha 8$  y G2, con estructuras de lipasas en conformaciones abiertas.

La presencia de estas estructuras hélices  $\alpha$  en la lipasa 2-17 podrían estar indicando que el fenómeno de activación interfacial está presente en esta enzima, y sería interesante realizar el modelo en base a un molde cuya estructura se encuentre cerrada, para estudiar el movimiento de estas dos hélices y conocer su posicionamiento cuando bloquean al sitio activo.

## **CONCLUSIONES GENERALES**

En el presente trabajo se logró realizar un modelo tridimensional de la estructura de la lipasa 2-17, que cumplió con los requerimientos básicos del programa Modeller. Este modelo fue bien evaluado por programas tipo Procheck, Verify 3D y VADAR. Particularmente para la evaluación con VADAR, un 61% de los aminoácidos mal evaluados correspondieron a aminoácidos presentes en *loops* de la proteína. El refinamiento del modelo realizado con Modeller fue enfocado particularmente en 10 *loops* que presentaron los principales problemas estructurales, y cuyo ordenamiento y evaluación final aumentó el número de aminoácidos con mala evaluación, pero cuyos rangos en los problemas detectados disminuyeron significativamente en comparación a los valores obtenidos en el modelo previo al proceso de refinamiento.

Para corroborar el modelo obtenido con la herramienta computacional Modeller se utilizaron los servidores Swiss Model e I-TASSER, cuyos modelos entregados lograron satisfacer las barreras evaluativas y cuyas estructuras fueron básicamente las determinadas con Modeller.

El modelo obtenido entonces para la lipasa estudiada consistió en una estructura tipo  $\alpha/\beta$  hidrolasa, compuesta por ocho hojas  $\beta$  centrales entre las cuales siete se encuentran ubicadas de forma paralela, mientras que solo una de manera antiparalela, formando una estructura semicilíndrica y cuyo dominio central está formado desde los residuos W126 hasta A415. La hoja  $\beta$  se encuentra rodeada por 15 estructuras secundarias, dentro de las cuales 11 constituyen el clásico plegamiento tipo  $\alpha$  hélices y 4 corresponden a pequeñas hélices con giros opuestos, llamadas estructuras G.

El alineamiento de la lipasa en estudio con la lipasa molde reveló tres posibles aminoácidos como principales candidatos para la triada catalítica: Serina 239, Aspartato 361 e Histidina 391, ubicados de manera canónica en el plegamiento tipo  $\alpha/\beta$  hidrolasa y cuya distribución espacial permite formar interacciones entre ellos para producir la hidrólisis del sustrato.

La composición del sitio activo de la proteína molde reveló una importante similitud con la composición del sitio activo de la lipasa 2-17. La densidad de aminoácidos del molde fue bastante mayor, presentando cerca de un 14% más de aminoácidos en una radio de 6Å que en la lipasa 2-17 y por consiguiente presentando un mayor número de enlaces. Esto podría estar implicado en las propiedades extremófilas de cada una de las enzimas. La comparación del sitio activo de la lipasa en estudio, con otra lipasa psicrófila reveló algunas diferencias en las composiciones de los aminoácidos pero una similitud en las densidades aminoacídicas presentadas en el sitio activo.

Con el objetivo de conocer la interacción de la lipasa en estudio y definir un sustrato preferente, se realizó *docking* molecular de los diferentes sustratos estudiados en el capítulo 3, y se seleccionaron 4 tipos de triacilgliceroles formados por los ácidos grasos Caproico (de 6 carbonos), Caprílico (de 8 carbonos), Cáprico (de 10 carbonos) y Palmítico (de 16 carbonos). Se realizó *docking* molecular ciego que

define un espacio de búsqueda suficientemente grande de modo de permitir la unión del ligando en cualquier lugar del receptor y *docking* forzado que define un espacio de búsqueda limitado al área de interacción observada en otros dominios de unión de las lipasas y a los aminoácidos seleccionados como posibles candidatos para la triada catalítica. Los resultados evidenciaron que la lipasa 2-17 tiene afinidad por sustrato de cadenas de 6 carbonos, mientras que su preferencia disminuye a medida que aumenta el número de carbonos en las cadenas aciladas.

Para corroborar la participación de los aminoácidos propuestos como miembros de la triada catalítica se realizaron cinco mutaciones *in silico* que correspondieron a los cambios de la Serina 239 por una Asparagina (S239N), del Aspartato 361 por una Alanina (D361A), una Asparagina (D361N) y un Glutamato (361E), y finalmente de la Histidina 391 por una Asparagina (H391N).

Mediante el modelamiento por homología de los mutantes, se estudiaron los cambios producidos por las mutaciones a nivel estructural. Se realizó un análisis de la disposición espacial de los aminoácidos y su relación con los aminoácidos del sitio activo. Para la mutación S239N no se observó interacción entre los aminoácidos implicados, posiblemente ya que la asparagina es un aminoácido más grande que además de no poder sustituir la función catalítica de la serina, provoca un pequeño cambio en el posicionamiento espacial tanto de la triada como del sustrato. Para el caso de la mutación D361A la lejanía y la naturaleza no reactiva de la alanina impidieron la formación de enlaces. La mutación D361E tampoco mostró interacciones entre los aminoácidos ya que si bien el glutamato tiene disponible un oxígeno reactivo, la nueva disposición de este aminoácido no permite la interacción de la triada. La misma situación fue posible apreciar en el caso de la variante D361N, donde la nueva disposición aminoacídica impidió la interacción entre la triada. Finalmente para la mutación H391N la falta del grupo imidazol y la activación de las cargas, además del aumento en las distancias atómicas influyeron directamente en la triada.

Para conocer la interacción de cada una de las variantes de la enzima en estudio, se procedió a realizar *docking* molecular para cada una de ellas con el sustrato preferente para esta lipasa, triacilglicerol con cadenas de 6 carbonos. En todos los casos, las energías de interacción de cada una de las enzimas mutantes fueron significativamente mayores a las obtenidas para la enzima nativa. A pesar de que los resultados anteriores mostraron una muy baja hidrólisis de los diferentes sustratos para cada una de las mutaciones, se procedió a realizar *docking* molecular de cada una de las mutaciones con los diferentes sustratos. En todos los casos, nuevamente las energías de interacción mostraron la falta de interacción entre los aminoácidos de las diferentes mutantes y los sustratos.

Con el objetivo de corroborar los resultados obtenidos *in silico*, se realizaron las mismas mutaciones propuestas pero esta vez mediante mutagénesis sitio-dirigida. Para cada uno de los casos se

estudiaron las cinéticas de reacción comparadas con la enzima nativa. Ninguna de las mutaciones evidenció actividad catalítica ensayada con el sustrato *p*-nitrofenil caprato.

Para descartar que la falta de actividad no estuviese dada por la falta de enzima en la muestra, se midieron las concentraciones de enzima, obteniéndose una variación máxima de tan solo un 18% para la mutación D361A, por lo que la falta de actividad se atribuyó a la ineficiencia catalítica de cada variante de la enzima.

Para descartar que los mutantes no hubiesen cambiado las temperaturas a las cuales se obtiene su máxima actividad, se realizaron ensayos de cada uno de los mutantes en un rango de temperaturas de 5°C a 35°C. En todos los casos las variantes de la enzima no exhibieron actividad catalítica. El mismo análisis fue realizado para la dependencia del pH. Se ensayaron cada una de las variantes en rangos de pH entre 4 y 11, pero no se observó actividad catalítica en ninguno de los clones mutantes. Finalmente se evaluó la posible variación en la especificidad de sustrato de los mutantes, debido principalmente a las modificaciones espaciales del sitio activo ya que éstas pueden influir en la selectividad por un determinado sustrato. Se realizaron las curvas de progreso con sustratos de cadenas de 2C, 6C, 8C, 10C y 16C y se ensayaron cada una de las enzimas variantes, pero no se observó actividad.

Esto demostró la importancia de estos tres aminoácidos en la catálisis enzimática, respaldando la hipótesis de que la posible triada catalítica de la lipasa 2-17 corresponda a la Serina 239, Aspartato 361 e Histidina 391.

Finalmente, para determinar si la lipasa 2-17 presenta el fenómeno de activación interfacial, se estudió la estructura tridimensional de esta enzima y se comparó con la estructura tridimensional de una lipasa que tuviese estructuras secundarias que actúen como tapa y cuya activación interfacial es conocida. Este análisis reveló dos posibles estructuras tipo  $\alpha$  hélices, G2 y  $\alpha$ 8, que podrían estar mediando la entrada del sustrato al sitio activo y por lo tanto que podrían estar evidenciando que la lipasa 2-17 presenta activación interfacial.



## **BIBLIOGRAFÍA**

1. Acevedo, J.P., Reyes F., Parra, L., Salazar, O., Andrews, B.A. and Asenjo, J.A. (2008) Cloning of complete genes for novel hydrolytic enzymes from Antarctic sea water bacteria by use of an improved genome walking technique. *Journal of Biotechnology*, 133: 277-286.
2. Acevedo, J.P., Rodríguez, V., Saavedra, M., Muñoz, M., Salazar, O., Asenjo, J.A. and Andrews, B. (2013) Cloning, expression and decoding of the cold adaptation of a new widely represented thermolabile subtilisin-like protease. *J. Applied Microbiology*, 114: 352-363.
3. Agarwal, P. K. (2006). Enzymes: An integrated view of structure, dynamics and function. *Microb. Cell Fact.* 5: 2.
4. Aghajari, N., Van Petegem, F., Villeret, V., Chessa, J. P., Gerday, C., Haser, R. and Van Beeumen, J. (2003). Crystal structures of a psychrophilic metalloprotease reveal new insights into catalysis by cold-adapted proteases. *Proteins*. 50: 636-647.
5. Alexandrov, V., Lehnert, U., Echols, N., Milburn, D., Engelman, D. and Gerstein, M. (2005). Normal modes for predicting protein motions: a comprehensive database assessment and associated web tool. *Prot. Sci.* 14: 633-643.
6. Arnold, K., Bordoli, L., Kopp, J. and Schwede, T. (2006). The SWISS-MODEL Workspace: A web-based environment for protein structure homology modelling. *Bioinformatics*. 22:195-201.
7. Arpigny, J. L. and Jaeger, K.E. (1999). Bacterial lipolytic enzymes: classification and properties. *Biochem. J.* 343:177-183.
8. Arpigny, J.L., Lamotte, J. and Gerday, C. (1997). Molecular adaptation to cold of an antarctic bacterial lipase. *J. Mol. Catal. B Enzym.* 3: 29-35.
9. Bae, E. and Phillips, G. N. Jr. (2004). Structures and analysis of highly homologous psychrophilic, mesophilic and thermophilic adenylate kinases. *J. Biol. Chem.* 279: 28202-28208.
10. Bahar, I. and Jernigan, R. L. (1997). Inter-residue potentials in globular proteins and the dominance of highly specific hydrophilic interactions at close separation. *J. Mol. Biol.* 266: 195-214.
11. Bai, Y. W., Sosnick, T.R., Mayne, L. and Englander, S. W. (1995). Protein-folding intermediates: native-state hydrogen-exchange. *Science*. 269: 192-197.
12. Baldwin, R. (1986). Temperature dependence of the hydrophobic interaction in protein folding. *PNAS*. 83: 8069-8072.
13. Bassegoda, A., Cesarini, S. and Diaz, P. (2012). Lipase improvement: goals and strategies. *Computational and Structural Biotechnology Journal*. 2:2-8.
14. Benkert, P., Tosatto, S.C. E. and Schomburg, D. (2008). QMEAN: A comprehensive scoring function for model quality assessment. *Proteins: Structure, Function, and Bioinformatics*. 71(1):261-277.
15. Benkovic, S. J. and Hammes-Schiffer, S. (2003). A perspective on enzyme catalysis. *Science*. 301: 1196-1202.
16. Bernstein, F.C., Koetzle, T.F., Williams, G.J.B., Meyer, E.F., Brice, M.D., Rodgers, J.R., Kennard, O., Shimanouchi, T. and Tasumi, M. (1977). The protein data bank: A computer-based archival file for macromolecular structures. *Journal of molecular biology*. 112(3): 535-542.
17. Biasini, M., Bienert, S., Waterhouse, A., Arnold, K., Studer, G., Schmidt, T., Kiefer, F., Gallo Cassarino, T., Bertoni, M., Bordoli, L. and Schwede, T. (2014). SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Research*. doi: 10.1093/nar/gku340.
18. Böde, C., Kovács, I., Szalay, M., Palotai, R., Korcsmáros, T. and Csermely, P. (2007). Network analysis of proteindynamics. *FEBS Letters*. 581: 2776-2782.
19. Bowie, J. U., Lüthy, R. and Eisenberg, D. (1991). A method to identify protein sequences that fold into a known three-dimensional structure. *Science*. 253(5016):164-70.
20. Brady, K., Brzozowski, A.M., Derewenda, Z.S., Dodson, E., Dodsom, G., Tolley, S., Turkcunburg, J.P., Christiansen, I., Høge-Jensen, B., Nørskov, L., Thim, K. and Mcnige, J. (1990). A sermeprotease triad forms the catalytic centre of a triacylglycerol lipase. *Nature*. 343: 767-770.

21. Bruice, T. C. and Benkovic, S. J. (2000). Chemical basis for enzyme catalysis. *Biochemistry*. 39: 6267-6274.
22. Brumlik, M. and Buckley, T. (1996). Identification of the Catalytic Triad of the Lipase/Acyltransferase from *Aeromonas hydrophila*. *Journal of Bacteriology*. 178: 2060-2064.
23. Brzozowski, A.M., Derewenda, U. and Derewenda, Z.S. (1990). A model for interfacial activation in lipases from the structure of a fungal lipase-inhibitor complex. *Nature*. 351: 491-494.
24. Byun, J. S., Rhee, J.K., Kim, N.D., Yoon, J.H., Kim D. H., Koh, E., Oh, J. W. and Cho, H. S. (2007). Crystal structure of hyperthermophilic esterase EstE1 and the relationship between its dimerization and thermostability properties. *BMC Structural Biology* 10: 7-47.
25. Cavicchioli, R., Saunders, N. and Thomas, T. (2002). Cold shock response in microorganisms. In *Extremophiles for the UNESCO Encyclopedia of Life Support Systems*. in press. Chapter within a broad compilation on extremophiles, specifically covering coldshock adaptation of all cellular components of Bacteria, Archaea and Eucarya.
26. Cavicchioli, R., Siddiqui, K. S., Andrews, D. and Sowers, K.R. (2002). Low-temperature extremophiles and their applications. *Curr. Opin. Biotechnol.* 13: 253-261.
27. Cavicchioli, R., Thomas, T., Lederberg, J., Alexander, M., Bloom, B.R., Hopwood, D., Hull, R., Iglewski, B.H., Laskin, A.I., Oliver, S.G., Schaechter, M. and Summers, W. C. (2000). Extremophiles. In *Encyclopedia of Microbiology*, edn 2. Broad coverage of extremophiles including sections on habitats and microorganisms. The biochemistry and physiology of adaptation of psychrophiles are also covered. *San Diego: Academic Press Inc.* 317-337.
28. Chen, V.B., Arendall, W.B., Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J. S. and Richardson, D.C. (2009). MolProbity: all-atom structure 79 validation for macromolecular crystallography. *Acta Crystallographica Section D: Biological Crystallography*. 66(1): 12-21.
29. Cheng, M., Angkawidjaja, C., Koga, Y. and Kanaya, S. (2012). Requirement of lid2 for interfacial activation of a family I.3 lipase with unique two lid structures. *The FEBS Journal*. 279: 3727-3737.
30. Cherry, J. R. and Fidantsef, A.L. (2003). Directed evolution of industrial enzymes: an update. *Curr. Opin. Biotech.* 14: 438-443.
31. Chothia, C. and Lesk, A. M. (1986). The relation between the divergence of sequence and structure in proteins. *The European Molecular Biology Organization*. 5: 823-826.
32. Clyde, S., Hutchison, A., Marshal, E., Shirley, P. and Smith, M. (1978) Mutagenesis at a Specific Position in a DNA Sequence. *The Journal of Biological Chemistry*. 253: 6551-6560.
33. Collins, T., Meuwis, M. A., Gerday, C. and Feller, G. (2003). Activity, stability and flexibility in glycosidases adapted to extreme thermal environments. *J. Mol. Biol.* 328: 419-428.
34. Corey, D. R. and Craik, C. S. (1992). An investigation into the minimum requirements for peptide hydrolysis by mutation of the catalytic triad of trypsin. *J. Am. Chem. Soc.* 114: 1784-1790.
35. Craik, C. S., Roczniak, S., Largman, C. and Rutter, W. J. (1987). The catalytic role of the active site aspartic acid in serine proteases. *Science*. 237: 909-913.
36. D'Amico, S., Claverie, P., Collins, T., Georlette, D., Gratia, E., Hoyoux, A., Meuwis, M. A., Feller, G. and Gerday, C. (2002). Molecular basis of cold adaptation. *Philos. Trans. R. Soc. London Ser. B*. 357: 917-925.
37. D'Amico, S., Marx, J.C., Gerday, C., and Feller, G. (2003). Activity-Stability Relationships in Extremophilic Enzymes. *The journal of biological chemistry*. 278: 7891-7896.
38. Daniel, R. M., Dunn, R. V., Finney, J. L. and Smith, J. C. (2003). The role of dynamics in enzyme activity. *Annu. Rev. Biophys. Biomol. Struct.* 32: 69-92.
39. Davis, I. W., Murray, L.W., Richardson, J.S. and Richardson, D.C. (2004). MOLPROBITY: structure validation and all-atom contact analysis for nucleic acids and their complexes. *Nucleic Acids Research*. 32: 615-619.
40. DeLano, W. L. (2002). The PyMOL Molecular Graphics System. DeLano Scientific, San Carlos, CA, USA.

41. Demirjian, D. C., Moris-Varas, F. and Cassidy, C.S. (2001). Enzymes from extremophiles. *Curr. Opin. Chem. Biol.* 5: 144-151.
42. De Simone G., Menchise V., Manco G., Mandrich L., Sorrentino N., Lang D., Rossi M. and Pedone C. (2001). The Crystal Structure of a Hyper-thermophilic Carboxylesterase from the Archaeon *Archaeoglobus fulgidus*. *J. Mol. Biol.* 314: 507-518.
43. Desnuelle, P., Sarda, L. y Ailhaud, G. (1960). Inhibition de la lipase pancreatique par le diethyl-p-nitrophenyl phosphate en emulsion. *Biochimica et Biophysica Acta.* 37, 570-571.
44. Dill, K. and Bromberg, S. (2003). Molecular driving forces: statistical thermodynamics in chemistry and biology. *Garland Science, New York.*
45. Dodson, G. and Verma, C. S. (2006). Protein flexibility: its role in structure and mechanism revealed by molecular simulations. *Cell. Mol. LifeSci.* 63: 207-219.
46. Dodson, G., and Wlodawer, A., (1998). Catalytic triads and their relatives. *Trends in Biochemical Sciences.* 23: 347-352.
47. Duhovny, D., Nussinov, R. and Wolfson, H. (2002) Efficient Unbound Docking of Rigid Molecules. *Springer-Verlag Berlin Heidelberg.* 185–200.
48. Duilio, A., Tutino, M. L. and Marino, G. (2004). Recombinant protein production in Antarctic Gram-negative bacteria. *Meth. Mol. Biol.* 267: 225-237.
49. Eisenmesser, E., Millet, O., Labeikovsky, W., Korzhnev, D., Wolf-Watz, M., Bosco, D., Skalicky, J., Kay, L. and Kern, D. (2005). Intrinsic dynamics of an enzyme underlies catalysis. *Nature.* 438: 117-121.
50. Epstein, C. J., Goldberger, R.F., Anfinsen, C. B. (1963). The genetic control of tertiary protein structure: studies with model systems. *Cold Spring Harbor Symposia on Quantitative Biology.* 28:439-449.
51. Eswar, N., Webb, B., Marti-Renom, M.A., Madhusudhan, M.S., Eramian, D., Shen, M. Y., Pieper, U. and Sali, A. (2007). Comparative protein structure modelling using Modeller. *Curr Protoc Protein Sci.* 50(2.9):1-2.9.
52. Ewing, T., Makino, S., Skillman, A. G. and Kuntz, I. D., (2001). Search strategies for automated molecular docking of flexible molecule databases. *Journal of Computer-Aided Molecular Design.* 15: 411–428.
53. Feller, G. (2003). Molecular adaptations to cold in psychrophilic enzymes. *Cell. Mol. LifeSci.* 60: 648-662.
54. Feller, G. (2007). Life at low temperatures: is disorder the driving force? *Extremophiles.* 11: 211-216.
55. Feller, G. and Gerday, C. (1997). Psychrophilic enzymes: molecular basis of cold adaptation. *Cell. Mol. Life Sci.* 53: 830-841.
56. Feller, G. and Gerday, C. (2003). Psychrophilic enzymes: hot topics in cold adaptation *Nat. Rev. Microbiol.* 1(3):200-208.
57. Feller, G., Arpigny, J. L., Narinx, E. and Gerday, Ch. (1997). Molecular Adaptations of Enzymes from Psychrophilic Organisms. *Camp. Biochem. Physiol.* 3: 495-499
58. Fenimore, P. W., Frauenfelder, H., McMahon, B. H. and Parak, F. G. (2002). Slaving: solvent fluctuations dominate protein dynamics and functions. *Proc. Natl. Acad. Sci. USA.* 99: 16047-16051.
59. Fields, P. (2001). Review: Protein function at thermal extremes: balancing stability and flexibility. *Comp. Biochem. Physiol.* 129: 417-431.
60. Fischer, M. and Pleiss, J. (2003). The lipase engineering database: a navigation and analysis tool for protein families. *Nucleic Acids Res.* 31:319-321.
61. Fiser, A., Gian Do, R. and Sali A. (2000). Modeling of loops in protein structures. *Protein Science.* 9: 1753-1773.
62. Fleming, P. and Richards, F. (2000). Protein packing: dependence on protein size, secondary structure and amino acid composition. *J. Mol. Biol.* 299: 487-498.

63. Fojan, P., Jonson, P.H., Petersen, M.T. and Petersen, S.B. (2000). What distinguishes an esterase from a lipase: a novel structural approach. *Biochimie* 82: 1033-1041.
64. Garcia-Viloca, M., Gao, J., Karplus, M. and Truhlar, D. G. (2004). How enzymes work: analysis by modern rate theory and computer simulations. *Science*. 303: 186-195.
65. Gerday, C., Aittaleb, M., Arpigny, J.L., Baise, E., Chessa J. P., Garsoux, G., Petrescu, I. and Feller, G. (1997). Psychrophilic enzymes: a thermodynamic challenge. *Biochim. Biophys. Acta* 1342: 119-131.
66. Gerday, C., Aittaleb, M., Bentahir, M., Chessa, J. P., Claverie, P., Collins, T., D'Amico, S., Dumont, J., Garsoux, G., Georgette, D., Hoyoux, A., Lonhienne, T., Meuwis, M. A. and Feller, G. (2000). Cold-adapted enzymes: from fundamentals to biotechnology. *Trends Biotechnol.* 18: 103-107.
67. Gerstein, M. and Krebs, W. (1998). A database of macromolecular motions. *Nucl. Acids Res.* 26: 4280-4290.
68. Gianese, G., Argos P. and Pascarella, S. (2001). Structural adaptation of enzymes to low temperatures. *Protein Eng.* 14: 141-148.
69. Giver, L., Gershenson, A., Freskgard, P. O. and Arnold, F. H. (1998). Directed evolution of a thermostable esterase. *Proc. Natl. Acad. Sci. USA* 95: 12809-12813.
70. Gomes, J. and Steiner, W. (2004). The biocatalytic potential of extremophiles and extremozymes. *Food Technol. Biotechnol.* 42: 223-235.
71. Gromiha, M. M. and Selvaraj, S. (2004). Inter-residue interactions in protein folding and stability. *Prog. Biophys. Mol. Biol.* 86: 235-277.
72. Guex, N. and Peitsch, M.C. (1997). SWISS-MODEL and the Swiss-Pdb Viewer: An environment for comparative protein modeling. *Electrophoresis.* 18(15): 2714-2723.
73. Guex, N., Diemand, A. and Peitsch, M. C. (1999). Protein modelling for all. *Trends Biochem. Sci.* 24: 364-367.
74. Guex, N., Peitsch, M. C., Schwede, T. (2009). SWISS-MODEL and Swiss PdbViewer. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective. *Electrophoresis.* 30(S1), S162-S173.
75. Gupta, R., Gupta, N. and Rathi, P. (2004). Bacterial lipases: an overview of production, purification and biochemical properties. *Appl. Microbiol. Biotechnol.* 64: 763-781.
76. Gunasekaran, K. and Nussinov, R. (2007). How different are structurally flexible and rigid binding sites? Sequence and structural features discriminating proteins that do and do not undergo conformational change upon ligand binding. *J. Mol. Biol.* 365: 257-273.
77. Halle, B. (2002). Flexibility and packing in proteins. *Proc. Natl. Acad. Sci. USA* 99: 1274-1279.
78. Hamelryck, T. and Manderick, B. (2003). PDB file parser and structure class implemented in Python. *Bioinformatics.* 19: 2308-2310.
79. Hei, D. J. and Clark, D.S. (1994). Pressure stabilization of proteins from extreme thermophiles. *Appl. Environ. Microbiol.* 60: 932-999.
80. Hide, A.W., Chan, L. and Li, W. (1992). Structure and evolution of the lipase superfamily. *Journal of Lipid Research.* 23: 167-178.
81. Hilton, S. and Buckley, J. T. (1991). Studies on the reaction mechanism of a microbial lipase/acyltransferase using chemical modification and site-directed mutagenesis. *J. Biol. Chem.* 266: 997-1000.
82. Hinsen, K. (2000). The molecular modelling tool kit: a new approach to molecular simulations. *J. Comp. Chem.* 21: 79-85.
83. Hough, D. W. and Danson, M.J. (1999). Extremozymes. *Curr. Opin. Chem. Biol.* 3: 39-46.
84. Hoyoux, A., Blaise, V., Collins, T., D'Amico, S., Gratia, E., Huston, A.L., Marx, J.C., Sonan, G., Zeng, Y., Feller, G. and Gerday, C. (2004). Extreme catalysts from low-temperature environments. *J. Biosci. Bioeng.* 98: 317-330.

85. Jaeger, K.E., Dijkstra, B. W. and Reetz, M.T. (1999). Bacterial biocatalysts: molecular biology, three-dimensional structures, and biotechnological applications of lipases. *Annu. Rev. Microbiol.* 53:315-351.
86. Jaeger, K. E. and Eggert, T. (2002). Lipases for biotechnology. *Curr. Opin. Biotechnol.* 13:390-397.
87. Jaenicke, R. (1991). Protein stability of preteins in extreme conditions. *Eur. J. Biochem.* 202: 715-728.
88. Jaenicke, R. (1998). The stability of proteins in extreme environments. *Curr. Opin. Struct. Biol.* 8:738-748.
89. Jaenicke, R. (2000). Stability and stabilization of globular proteins in solution. *J. Biotechnol.* 79: 193-203.
90. Jaenicke, R. and Závodszky, P. (1990). Proteins under extreme physical conditions. *FEBS Lett.* 268(2): 344-349.
91. Jager, S., Demleitner, G. and Gotz, F. (1992). Lipase of *Staphylococcus hyicus*: Analysis of the catalytic triad by site-directed mutagenesis. *FEMS Microbiology Letters.* 100: 249-254.
92. Ji, L., Xiaoling, T., Hongwei, Y. (2010). Prediction of the enantioselectivity of lipases and esterases by molecular docking method with modified force field parameters. *Biotechnology and Bioengineering.* 105: 687-696.
93. Jordan, F. and Polgár, L. (1981). Proton nuclear magnetic resonance evidence for the absence of a stable hydrogen bond between the active site aspartate and histidine residues of native subtilisins and for its presence in thiolsubtilisins. *Biochemistry* 20:6366-6370.
94. Juhl, P.B., Trodler, P., Tyagi, S. and Pleiss, J. (2009). Modelling substrate specificity and enantioselectivity for lipases and esterases by substrate-imprinted docking. *BMC Structural Biology.* 9:39-56.
95. Jung, S.K., Jeong, D. G., Lee, M. S., Lee, J. K., Kim, H. K., Ryu, S. E., Park, B. C., Kim, J. H., and Kim, S. J. (2008). Structural basis for the cold adaptation of psychrophilic M37 lipase from *Photobacterium lipolyticum*. *Wiley InterScience.* 71:476-484.
96. Kanaya, S., Koyanagi, T. and Nanaya, E. (1998). An esterase from *Escherichia coli* with a sequence similarity to hormone-sensitive lipase. *Biochem. J.* 332:75-80.
97. Kaplan, W. and Littlejohn, T.G. (2001). Swiss-PDB Viewer (Deep View). *Briefings in bioinformatics.* 2(2): 195.
98. Karplus, M. and Kuriyan, J. (2005). Molecular dynamics and protein function. *Proc. Natl. Acad. Sci. USA.* 102: 6679-6685.
99. Kiefer, F., Arnold, K., Künzli, M., Bordoli, L. and Schwede, T. (2009). The SWISS-MODEL Repository and associated resources. *Nucleic Acids Research.* 37: D387-D392.
100. Kim, K.K., Song, H.K., Shin, D.H., Hwang, K.Y. and Suh, S.W. (1997). The crystal structure of a triacylglycerol lipase from *Pseudomonas cepacia* reveals a highly open conformation in the absence of a bound inhibitor. *Structure.* 5:173-185.
101. Korman, T. and Bowie, J. (2012). Crystal Structure of *Proteus mirabilis* Lipase, a Novel Lipase from the Proteus/Psychrophilic Subfamily of Lipase Family I.1. *PLOS one.* 7:1-8.
102. Kumar, S., Ma, B., Tsai, C. J., Sinha, N. and Nussinov, R. (2000). Folding and binding cascades: Dynamic landscapes and population shifts. *Prot. Sci.* 9: 10-19.
103. Kumar, S., Wolfson, H. and Nussinov, R. (2001). Protein flexibility and electrostatic interactions. *IBM J. Res and Dev.* 45: 499-512.
104. Lonhienne, T., Gerday, C. and Feller, G. (2000). Psychrophilic enzymes: revisiting the thermodynamic parameters of activation may explain local flexibility. *Biochim. Biophys. Acta.* 1543: 1-10.
105. Lonhienne, T., Zoidakis, J., Vorgias, C. E., Feller, G., Gerday, C. and Bouriotis, V. (2001). Modular structure, local flexibility and cold-activity of a novel chitinase from a psychrophilic antarctic bacterium. *J. Mol. Biol.* 310: 291-297.

106. Lopez, C. F., Darst, R. K. and Rossky, P. J. (2007). Mechanistic elements of protein cold denaturation. *J. Phys. Chem. B*.112: 5961-5967.
107. López-Gallego, F., Abian, O., Guisán, J.M. (2012). Altering the Interfacial Activation Mechanism of a Lipase by Solid- Phase Selective Chemical Modification. *Biochemistry*. 51: 7028–7036.
108. Luo, Y., Zheng, Y., Jiang, Z., Ma, Y. and Wei, D. (2006). A novel psychrophilic lipase from *Pseudomonas fluorescens* with unique property in chiral resolution and biodiesel production via transesterification. *Appl. Microbiol. Biotechnol.* 73: 349-355.
109. Lüthy, R., Bowie, J. U., Eisenberg D. (1992). Assessment of protein models with three-dimensional profiles. *Nature*. 356(6364):83-5.
110. Ma, B., Kumar, S., Tsai, C. J. and Nussinov, R. (1999). Folding funnels and binding mechanisms. *Protein Eng.* 12: 713-720.
111. Margesin, R., Feller, G., Gerday, C. and Russell, N. (2002). Cold-adapted microorganisms: adaptation strategies and biotechnological potential. In Bitton, G. (ed.), *The encyclopedia of environmental microbiology*. John Wiley and Sons, New York, 871-885.
112. Marhuenda-Egea, F. C. and Bonete, M.J. (2002). Extreme halophilic enzymes in organic solvents. *Curr. Opin. Biotechnol.* 13: 385-389.
113. Martí-Renom, M., Stuart, A., Fiser, A., Sánchez, R., Melo, F. and Sali, A., (2000). Comparative protein structure modeling of genes and genomes. *Annual Review of Biophysics and Biomolecular Structure*.29: 291-325.
114. Marx, J. C., Collins, S., D'Amico, G., Feller, C. and Gerday, C. (2007). Cold-adapted enzymes from marine antarctic microorganisms. *Mar. Biotech.* 9: 293-304.
115. Matthews, B. W., Sigler, P. B., Henderson, R. and Blow, D. M. (1967). Three-dimensional structure of tosyl- $\alpha$ -chymotrypsin. *Nature*.214: 652–655.
116. Miyazaki, K., Wintrode, P. L., Grayling, R. A., Rubingh, D. N. and Arnold, F. H. (2000). Directed evolution study of temperature adaptation in a psychrophilic enzyme. *J. Mol. Biol.* 297: 1015-1026.
117. Morris, A.L., Macarthur, M.W., Hutchinson, E. G. and Thornton, J.M. (1992). Stereochemical quality of protein structure coordinates. *Proteins: Structure, Function, and Bioinformatics*. 12(4): 345-364.
118. Mulder, F. A. A., Mittermaier, A., Hon, B., Kahlquist, F. W. and Kay, L. E. (2001). Studying excited states of proteins by NMR spectroscopy. *Nature Struct. Biol.* 8: 932-935.
119. Nardini, M. and Dijkstra, B.W. (1999).  $\alpha/\beta$ hydrolase fold enzymes: the family keeps growing. *Curr. Opin. Struc. Biol.* 9:732-737.
120. Neves-Petersen, M.T., Fojan, P. y Petersen, S.B. (2001). How do lipases and esterases work: the electrostatic contribution. *Jurnal of Biotechnology*. 85: 115-147.
121. Panaiotov, I. and Verger, R. (2000). Enzymatic reactions at interfaces: Interfacial and temporal organization of enzymatic lipolysis. *Physical Chemistry of Biological Interfaces*. 359- 440.
122. Papaleo, E., Riccardi, L., Villa, C., Fantucci, P. and De Gioia, L. (2006). Flexibility and enzymatic cold-adaptation: A comparative molecular dynamics investigation of the elastase family. *Biochim. Biophys. Acta*.1764: 1397-1406.
123. Papaleo, E., Olufsen, M., De Gioia, L. and Brandsdal, B. O. (2007). Optimization of electrostatics as a strategy for cold-adaptation: a case study of cold- and warm-active elastases. *J. Mol. Graph. Mod.* 26: 93-103.
124. Parra, L., Reyes, F., Acevedo, J.P., Salazar, O., Andrews, B. and Asenjo, J. (2007). Cloning and fusion expression of a cold-active lipase from marine antarctic origin. *Enzyme and Microbial Technology*. 42: 371-377.
125. Pineda, J. R. E. T. and Schwartz, S. D. (2006). Protein dynamics and catalysis: the problems of transition state theory and the subtlety of dynamic control. *Phil. Trans. R. Soc. B* 361: 1433-1438.
126. Pleiss, J., Fischer, M. and Schmid, R.D. (1998). Anatomy of lipase binding sites: the scissile fatty acid binding sites. *Chem. Phys. Lipids*. 93(1-2):67- 80.

127. Pleiss, J., Fischer, M., Peiker, M., Thiele, C. and Schmid, R.D. (2000). Lipase engineering database. Understanding and exploiting sequence-structure-function relationships. *J. Mol. Catal. B Enzym.* 10:491-508.
128. Polgár, L. (2005). The catalytic triad of serine peptidases. *CMLS Cellular and Molecular Life Sciences.* 62: 2161-2172.
129. Rose, G., Fleming, P., Banavar, Jayanth and Maritan, A. (2006). A backbone-based theory of protein folding. *Proc. Natl. Acad. Sci. USA.* 103: 16623-16633.
130. Rost, B. (1999). Twilight zone of protein sequence alignments. *Protein Engineering Design and Selection.* 12: 85-94.
131. Rothschild, L. J. and Mancinelli, R.L. (2001). Life in extreme environments. *Nature.* 409: 1092-1101.
132. Roy A., Kucukural A., and Zhang Y. (2010). I-TASSER: a unified platform for automated protein structure and function prediction. *Nature Protocols.* 5: 725-738.
133. Roy A., Kucukural A., and Zhang Y. (2012). COFACTOR: An accurate comparative algorithm for structure-based protein function annotation. *Nucleic Acids Research.* 40: 471-477.
134. Russell, N. J. and Hamamoto, T. (1998). Psychrophiles. In *Extremophiles: Microbial Life in Extreme Environments.* Edited by Horikoshi, K., Grant and W. D. New York: *Wiley-Liss.* 25-45.
135. Russell, R. J., Gerike, U., Danson, M. J., Hough, D. W. and Taylor, G. L. (1998). Structural adaptations of the cold-active citrate synthase from an antarctic bacterium. *Structure.* 6: 351-361.
136. Sali, A. and Blundell, T.L. (1994). Comparative protein modelling by satisfaction of spatial restraints. *Protein Structure by Distance Analysis.* 64-86.
137. Sánchez, R. and Sali, A. (2000). Comparative protein structure modeling. *Protein Structure Prediction: Methods and Protocols, Human Press, New Jersey.*
138. Sambrook, J. and Russell, D.W. (2001). *Molecular cloning: a laboratory manual.* New York: *Cold Spring Harbor Laboratory Press.* ISBN 978-0879695774
139. Sánchez, R. and Sali, A. (2000). Comparative protein structure modeling. *Protein Structure Prediction: Methods and Protocols, Humana Press, New Jersey.*
140. Sander, C. and Schneider, R. (1991). Database of homology-derived protein structures and the structural meaning of sequence alignment. *Proteins.* 9: 56-68.
141. Sarda, H. and Desnuelle, P. (1958). Action de la lipase pancréatique sur les esters en émulsion. *Biochim. Biophys.* 30: 513-521.
142. Schiraldi, C. and De Rosa, M. (2002). The production of biocatalysts and biomolecules from extremophiles. *Trends Biotechnol.* 20: 515-521.
143. Schrag, J., Li, Y., Cygler, M., Lan, D., Burgdorf, T., Hecht, H., Schmid, R., Schomburg, D., Rydel, T., Oliver, J., Strickland, L., Dunaway, M., Larson, S., Day, J. and McPherson, A. (1996). The open conformation of a Pseudomonas lipase. *Structure.* 5: 187-202.
144. Schwede, T., Kopp, J., Guex, N. and Peitsch, M.C. (2003). SWISS-MODEL: an automated protein homology-modeling server. *Nucleic Acids Research.* 31: 3381-3385.
145. Sellek, G. A. and Chaudhuri, J.B. (1999). Biocatalysis in organic media using enzymes from extremophiles. *Enzym. Microb. Tech.* 25: 471-482.
146. Shen, M. and Sali, A. (2006). Statistical potential for assessment and prediction of protein structures. *Protein and Science.* 15: 2507-2524.
147. Shi, J., Blundell, T. and Mizuguchi, K. (2001). FUGUE: Sequence-structure homology recognition using environment-specific substitution tables and structure-dependent gap penalties. *Journal of Molecular Biology.* 310: 243-257.
148. Siddiqui, K. S. and Cavicchioli, R. (2006). Cold-adapted enzymes. *Annu. Rev. Biochem.* 75: 403-433.
149. Stobiecka, A., Wysockia, S. y Brzozowski, A.M. (1998). Fluorescence study of fungal lipase from *Humicola lanuginosa.* *Journal of Photochemistry and Photobiology.* 45: 95-102.



- 150.Svendsen, A. (2000). Lipase protein engineering. *Biochimica et Biophysica Acta*. 1543: 223-238.
- 151.Svingor, A., Kardos, J., Hajdu, I., Nemeth, A. and Zavodszky, P. (2001). A better enzyme to cope with cold. Comparative flexibility studies on psychrotrophic, mesophilic, and thermophilic PMDHs. *J. Biol. Chem.* 276: 28121-28125.
- 152.Van den Burg, B. (2003). Extremophiles as a source for novel enzymes. *Curr. Opin. Microbiol.* 6: 213-218.
- 153.Van Tilbeurgh, H., Egloff, M.P., Martinez Rugani, N., Verger, R. and Cambillau C. (1993). Interfacial activation of the lipase procolipase complex by mixed micelles revealed by X-ray crystallography. *Nature*. 362: 814-820.
- 154.Villeneuve, P., Muderhwa, J.M., Graille, J. and Haas, M. (2000). Customizing lipases for biocatalysis: a survey of chemical, physical and molecular biological approaches. *J. Molec. Catalysis B: Enzymatic*. 9:113-148.
- 155.Warshel, A. (1998). Electrostatic origin of the catalytic power of enzymes and the role of preorganized active sites. *J. Biol. Chem.* 273: 27035-27038.
- 156.Willard, I., Ranjan, A., Zhang, H., Monzavi, H., Boyko, R.F., Sykes, B.D. and Wishart, D.S. (2003). VADAR: a web server for quantitative evaluation of protein structure quality. *Nucleic Acids Research*. 31(13): 3316-3319.
- 157.Winkler, F.K., D'Arcy, A. and Hunziker, W. (1990). Structure of human pancreatic lipase. *Nature*. 343: 771-774.
- 158.Wintrode, P. L. and Arnold, F. H. (2000). Temperature adaptation of enzymes: lessons from laboratory evolution. *Adv. Protein Chem.* 55: 161-225.
- 159.Wooll, J. O., Wrabl, J. O. and Hilser, V. J. (2000). Ensemble modulation as an origin of denaturant-independent hydrogen exchange in proteins. *J. Mol. Biol.* 301: 247-256.
- 160.Zhang, Y. (2008). I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics*. 9: 40.
- 161.Zaccai, G. (2000). How soft is a protein? A protein dynamics force constant measured by neutron scattering. *Science*. 288: 1604-1607.
- 162.Zavodszky, P., Kardos, J., Svingor, A. and Petsko, G. (1998). Adjustment of conformational flexibility is a key event in the thermal adaptation of proteins. *Proc. Natl. Acad. Sci. USA*. 95: 7406-7411.
- 163.Zecchinon, L., Claverie, P., Collins, T., D'Amico, S., Delille, D., Feller, G., Georlette, D., Gratia, E., Hoyoux, A., Meuwis, M. A., Sonan, G. and Gerday, C. (2001). Did psychrophilic enzymes really win the challenge? *Extremophiles*. 5: 313-321.

## **ANEXOS**

## 8.1 ABREVIACIONES UTILIZADAS

1Kpb	Marcador de peso molecular 1Kbp
100bp	Marcador de peso molecular 100bp
Amp	Ampicilina
Aa	Aminoácidos
BSA	Seroalbúmina de bovino
CIBYB	Centro de Ingeniería Bioquímica y Biotecnología
dNTPs	Deoxiribonucleótidos trifosfato
DO	Densidad óptica
EDTA	Etilendiaminotetraacetato ácido
$g$	Fuerza de gravedad
g	Gramo
Ori	Origen de Replicación
HSL	Lipasa sensible a hormonas
IPTG	Isopropil tio- $\beta$ -D-galactósido
kDa	Kilodalton
LB	Medio Luria-Bertani
M	Molar
MBP	Proteína de unión a Maltosa
MCS	Sitio Múltiple de Clonamiento
min	Minutos
mL	Mililitros
ORF	Marco de lectura abierto
PAGE	Electroforesis en gel de poliacrilamida.
pb	Pares de bases
PCR	Reacción en cadena de la polimerasa
<i>p</i> -NFC	Para-nitrofenil caprato
rpm	Revoluciones por minuto
SDS	Dodecil sulfato de sodio
TEMED	N, N, N', N'-tetrametiletildiamina
Tris	Tris-(hidroximetil)-aminoetano
U	Unidad enzimática
$\mu$ F	Microfaradios
UV	Radiación ultravioleta
V	Voltio
X-GAL	5-bromo-4-cloro-3-indolil- $\beta$ -D-galactopiranosido

## 8.2 MODELAMIENTO DE LA LIPASA 2-17

### 8.2.1 Secuencia lipasa 1jji

>1JJI

MLDMPIDPVYYQLAEYFDSLKFDQFSSAREYREAINRIYEERNRQLSQHERVERVEDRTIKGRNGDIRVRVYQQKPDSP  
VLVYYHGGGFVICSIESHDALCRRRIARLSNSTVVSVDYRLAPEHKFPAAVYDCYDATKWVAENAEELRIDPSKIFVGGDSA  
GGNLAAAVSIMARSDGEDFIKHQILIYPVVNFVAPTPSLLEFGEGWLWDQKIMSWFSEQYFSREEDKFNPLASVIFADLE  
NLPPALIITAEDPLRDEGEVFGQMLRRAGVEASIVRYRGVLHGFINYYPVLKAARDAINQIAALLVFD

## 8.2.2 Alineamiento de las secuencias entregado por el archivo build\_profile.ali:

>P1;QUERY

sequence:QUERY: 0: : 0: :::-1.00:-1.00

MPILPVPSLNALLTKTLKTVKKTAANKVHQHHILHHTLKGLDNLPTPLLERVNHQLKASTAEQYPLADAHRLRLIL  
AVSNKLRPLALDKLPNLRQKFGTDAVSLQAPSVWQQNAEASGNSENAVSWQDKTIANADGGDMTVRCYQQSKQN  
NEGKSTDEVAMLFFHGGGFCIGDIDTHHEFCHTVCAQTGWAVVSVDYRMAPEYPAPTALKDCLSAYAWLTEHSQS  
LGALPSRIVLSGDSAGGCLAALVAQQVIKPIDALWQDNNQATETDKKANDTFKKS LADLPRPLAQLPLYPVTDYE  
AEYPSWELYGEGLLLDHNDANVFNTAYTQHSGLMQSHPLISVMHGDNAHLCPSYVVVAELDILRDEGLAYAELLO  
KEGVKVEYTYTVLGAPHGFINLMSVHQGLGNQNTNYIIKSFACFVQNLTTSEGDEPNL\*

>P1;1ea5A

structure:1ea5A: 18: : 210: :::-1.00:-1.00

-----PVLSSHISAFI  
GI-----PFAEPPVGNMR--FRRPEPKKPWSGVW----NASTYPNNCQQYVDEQFPGFSGSEMWNPNREMSDS  
PRPKST--TVMVWIYGGGFYSGS-STLDVYNGKYLAYTEEVVVSLSYRVSQEAPGNVGLLDQRMALQVVDNIQF  
FGGDPKTVTIFGESAGG---ASVGMHILSP-----  
-----\*  
-----\*

>P1;2bce

structure:2bce: 51: : 200: :::-1.00:-1.00

-----WQGTKAKSFKKQATLTQDSTYGNEDCLYLNIWVPQGRKE-  
---VSHDLPVMIWIYGGAFIMG-LSNYLYDGEEIATRGNVIVVTFNYRVGNLPGNYGLWDQHMAIAWVKRNIEA  
FGGDPDQITLFGESAGG---ASVSLQTLSP-----  
-----\*  
-----\*

>P1;1qe3A

structure:1qe3A: 90: : 219: :::-1.00:-1.00

-----MVWIHGGAFYLGAGSEPLYDGSKLAAQGEVIVVTLNYRLGPAYSDNLGLLDQAAALKWVRENISA  
FGGDPDNVTVFGESAGGSIAALLAMPAKGLKAIMESGASRTMTKEQAAST-----  
-----\*  
-----\*

>P1;1u4nA

structure:1u4nA: 55: : 291: :::-1.00:-1.00

-----  
-----GRTLKVRMYRP-----  
-EGVEPPYPALVYYHGGGWVVGDLETHDPVCR-VLAKDGRAVVSVDYRLAPEHKFPAAVEDAYDALQWIAERAAD  
FHLDPARIAVGGDSAGGNLAAV-----TSILAKERGGPALA-----FQLLIYPSTGYD  
PAHPSIEENAEGYLLTGGMSLWFLDQYL-NSLEELTHPWFSVLYPDLSSGLPPAYIATAQYDPLRDVGKLYAEALN  
KAGVKVEIENFEDLIHGFAQFYSLSPG-----\*

>P1;1f6wA

structure:1f6wA: 51: : 207: :::-1.00:-1.00

-----  
-----WQGTKLAKNFKKQATITQDSTYGDEDCLYLNIVWPQGRKQ-  
---VSRDLPVMIWIYGGAFLMGSLNNYLYDGEEIATRGNVIVVTFNYRVGNLPGNYGLRDQHMAIAWVKRNIAA  
FGGDPDNITLFGESAGG---ASVSLQTLSP-----  
-----\*

>P1;1jkmA

structure:1jkmA: 41: : 327: :::-1.00:-1.00

-----  
-----ASDDLPTVLAAVGASHDGFQA--VYDSIALDLPTDRDDVETSTETILGVDGNEITLHVFRPA---  
--GVEGVLPLVYTHGGGMTILTDDVHRRWC-TDLAAAGSVVVMVDFRNAGHHPFSGVEDCLA AVLWVDEHRES  
LGL--SGVVVQGESGGNLATLLAKRRGR-LDAI--DGVYASIPYISGGYAW-----DHERRLTELPSLVEND--  
-----GYFIENGMALLVRAYDEHAEDPIAWPYFA-SEDELRLPPFVAVNELDPLRDEGIAFARRLA  
RAGVDVAARVNIGLVHG-----\*

>P1;1k4yA

structure:1k4yA: 47: : 202: :::-1.00:-1.00

-----  
-----QPAESWSHVKNNTSYPPMCSQDAVSGHMLSELFTNRKEN  
IPLKFSEDCVMVWIHGGGLMVGASTYDGL--ALSAHENVVVVTIQYRLGDEHSRGNHLDQVAALRWVQDNIAN  
FGGDPGSVTIFGESAGG-----  
-----\*

>P1;1lz1A

structure:1lz1A: 3: : 290: :::-1.00:-1.00

-----FPTLDP ELAAAL  
TMLPKVD----FADLPNARATY--DALI-----GAMLADLSFDGVSLRELSAPGLD-GDPEVKIRFVTPDN  
TAGPVP---VLLWIHGGGFAIGTAESSDFCVEVARELGFAVANVEYRLAPETTFPGPVND CYAALLYIHAAHEE  
LGIDPSRIAVGGQSAGGGLAAKARDEGVVPVAFQFL---EIPELDDRLETVSMTNFVDTPRPNAIL-----  
----SWKYYGE-----SYSG--PEDPDVSIYAADLTGLPPTYLSTMELDPLRDEGIEYALRLL  
QAGVSVELHSFPGTFHG-----\*

>P1;1mx1A

structure:1mx1A: 49: : 217: :::-1.00:-1.00

-----  
-----QPAEPWSFVKNATSYPPMCTQDPKAGQLLSELFTRNKEN  
IPLKLS EDCVMVWIHGGGLMVGAASTYDGL--ALAAHENVVVVTIQYRLGDEHSRGNHGLDQVAALRWVQDN IAS  
FGGNPGSVTIFGESAGGSVSVLVLSPLAK-----  
-----  
-----\*

>P1;1p0iA

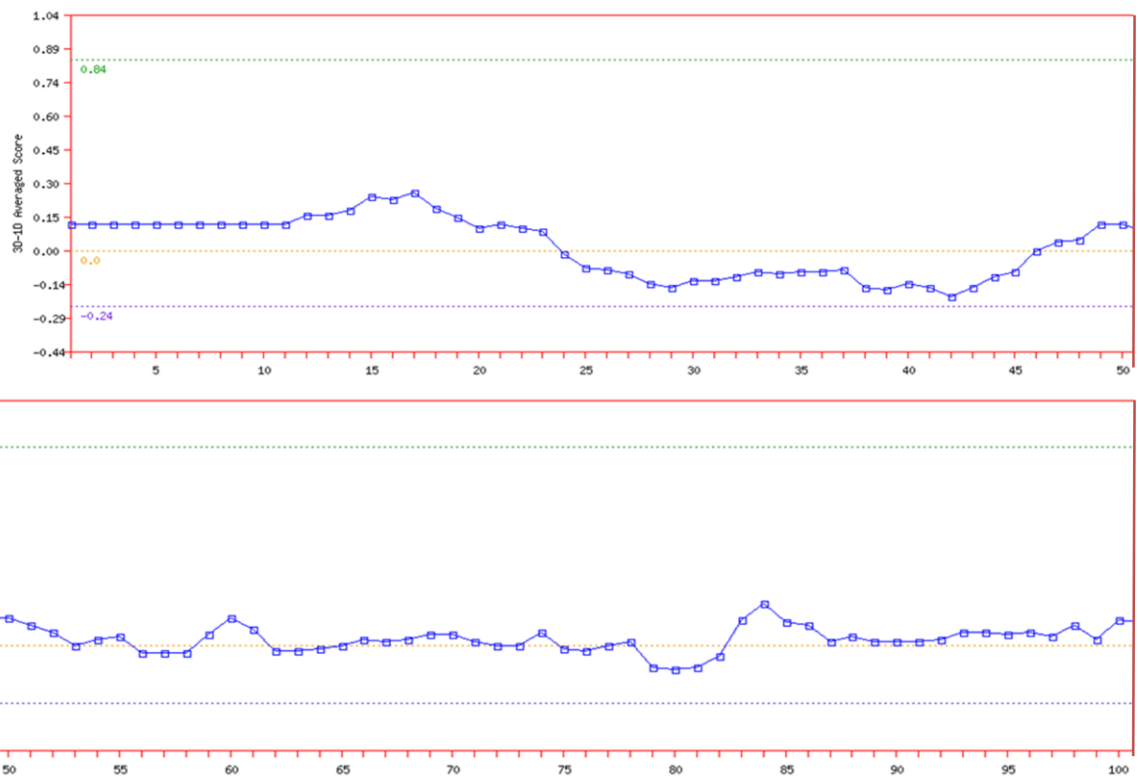
structure:1p0iA: 100: : 208: :::-1.00:-1.00

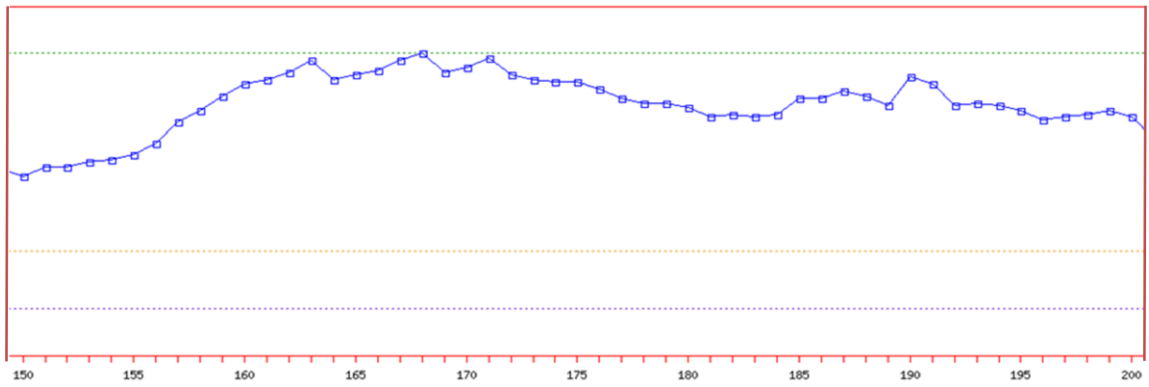
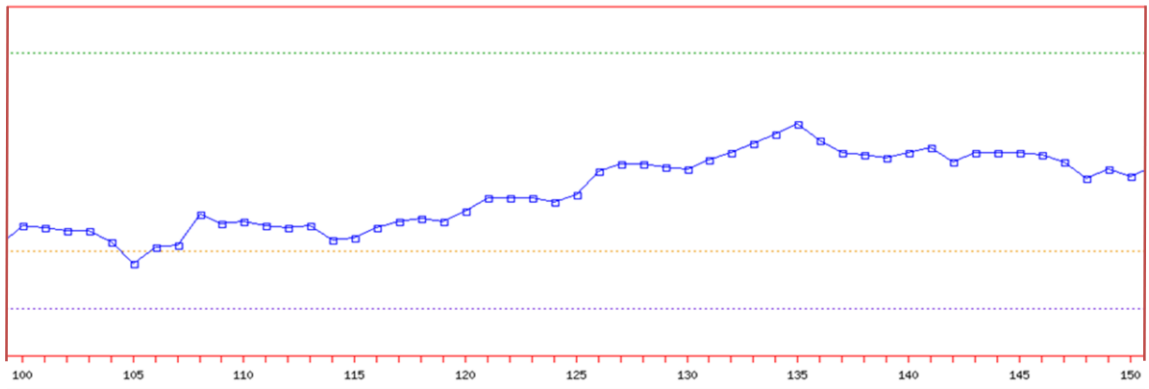
-----  
-----  
---KPKNATVLIWIYGGGFQTGTSSLHVYDGKFLARVERVIVVSMNYRVGPEAPGNMGLFDQQLALQWVQKNIAA  
FGGNPKSVTLFGESAG---AASVSLHLLSP-----  
-----  
-----\*

### 8.2.3 Valores de corte entregados por el programa FUGUE.

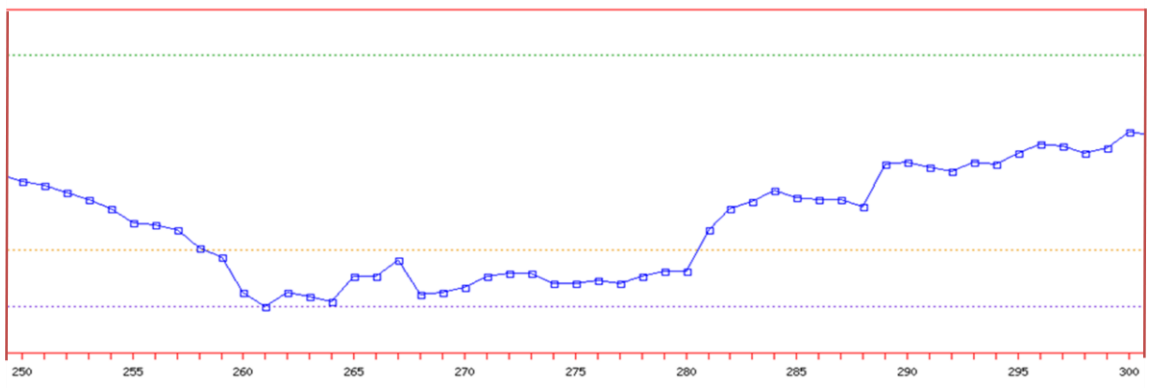
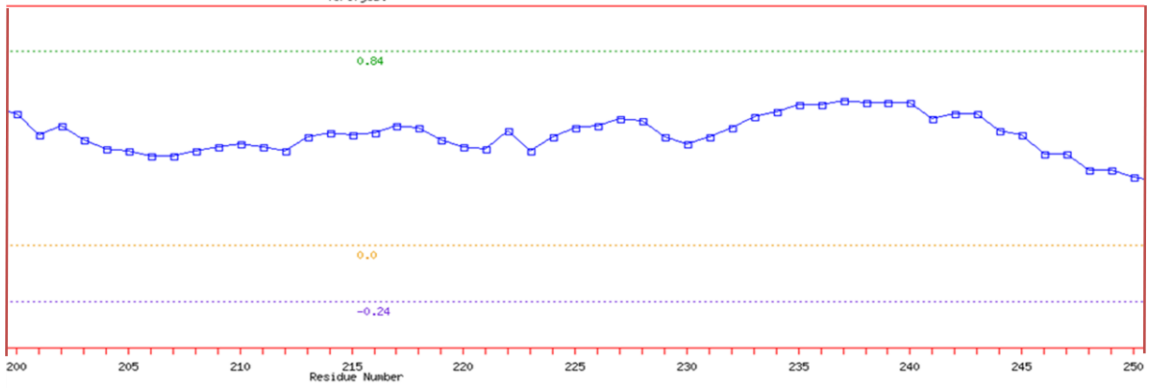
Recommended cutoff	ZSCORE	CONFIDENCE
		>= 6.0
Other cutoff	>= 4.0	LIKELY 95% confidence
Other cutoff	>= 3.5	MARGINAL 90% confidence
Other cutoff	>= 2.0	GUESS 50% confidence
Other cutoff	< 2.0	UNCERTAIN

### 8.2.4 Evaluación del modelo utilizando Verify 3D

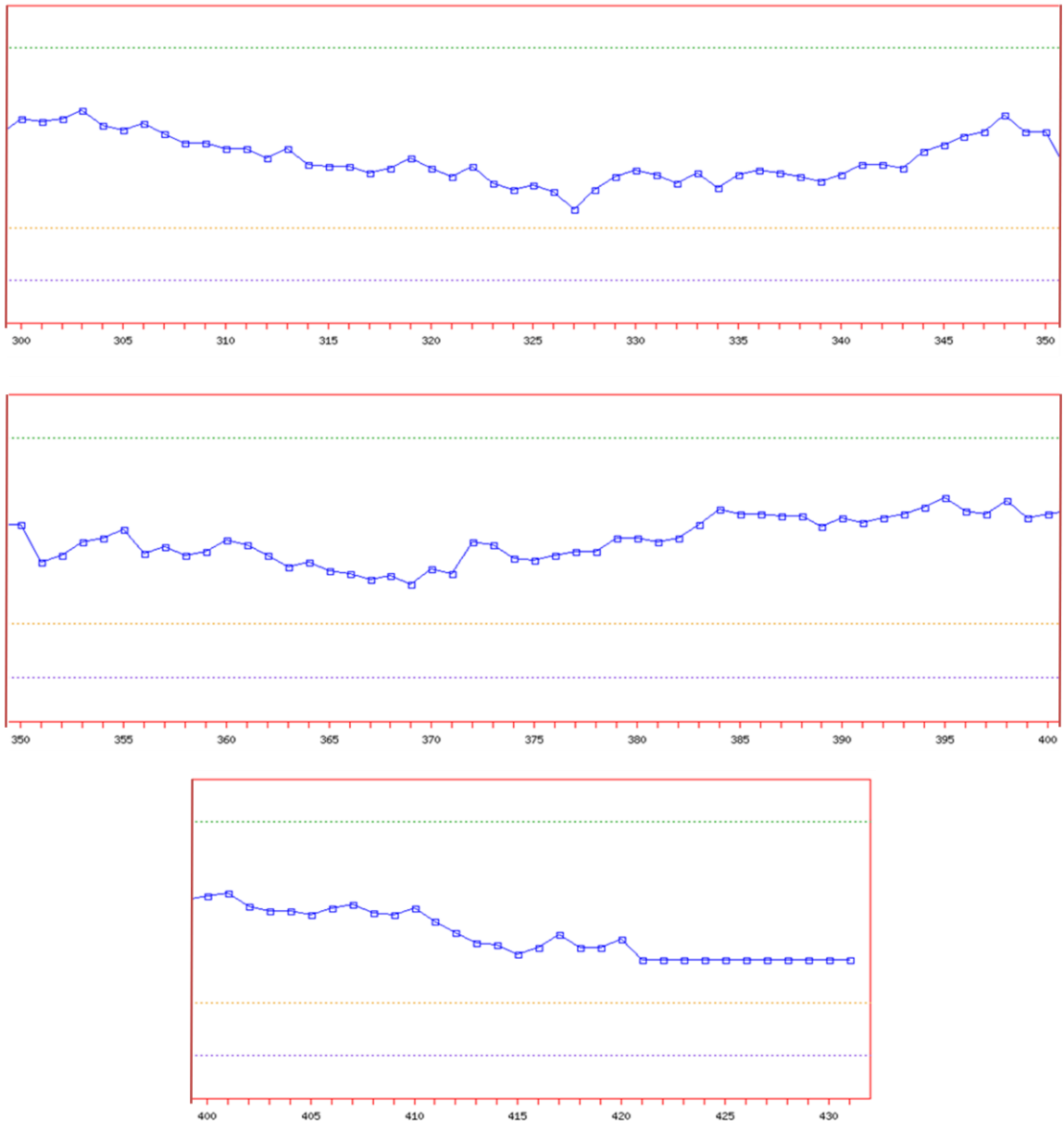




Verify3D:







**Figura 69: Perfil de valores entregados por el programa Verify 3D.** La línea azul corresponde a todos los aminoácidos presentes en el modelo 3, la línea amarilla corresponde al valor 0.0, la verde y la gris corresponden al mayor y al menor valor establecido, respectivamente. En el eje “y” se observa la relación promedio 3D-1D y en el eje “x” el número de cada aminoácido ordenado en la secuencia.

## 8.2.5 Refinamiento de los *loops* de la lipasa 2-17

### 8.2.5.1 Archivo de entrada

```
# Loop refinement of an existing model
from modeller import *
from modeller.automodel import *

log.verbose()
env = environ()

# directories for input atom files
env.io.atom_files_directory = './:../atom_files'

# Create a new class based on 'loopmodel' so that we can redefine
# select_loop_atoms (necessary)
class MyLoop(loopmodel):
    # This routine picks the residues to be refined by loop modeling
    def select_loop_atoms(self):
        # 10 residue insertion
        return selection(self.residue_range('259', '288'))

m = MyLoop(env,
           inimodel='Query.B999990003.pdb', # initial model of the target
           sequence='Query')              # code of the target

m.loop.starting_model= 1                 # index of the first loop model
m.loop.ending_model  = 10                # index of the last loop model
m.loop.md_level = refine.slow # loop refinement method; this yields
                              # models quickly but of low quality;
                              # use refine.slow for better models

m.make()
```

### 8.2.5.2 Resultados Refinamiento

#### Aminoácidos: 259-288

Filename	molpdf
Query.BL00010001.pdb	58.26345
Query.BL00020001.pdb	78.38163
Query.BL00030001.pdb	62.06118
Query.BL00040001.pdb	4860.33398
Query.BL00050001.pdb	6966.42676
Query.BL00060001.pdb	5383.28516
Query.BL00070001.pdb	81.55203
Query.BL00080001.pdb	1043.74390
Query.BL00090001.pdb	3490.44653
Query.BL00100001.pdb	145.22685

**Aminoácidos: 89 – 124**

Filename	molpdf
Query.BL00010001.pdb	94.55986
Query.BL00020001.pdb	54.32837
Query.BL00030001.pdb	84.46768
Query.BL00040001.pdb	2389.53931
Query.BL00050001.pdb	105.27953
Query.BL00060001.pdb	116.30473
Query.BL00070001.pdb	70.38745
Query.BL00080001.pdb	81.10049
Query.BL00090001.pdb	5785.16553
Query.BL00100001.pdb	3011.23340

**Aminoácidos: 20 – 60**

Filename	molpdf
Query.BL00010001.pdb	5948.20264
Query.BL00020001.pdb	84.42596
Query.BL00030001.pdb	67.30663
Query.BL00040001.pdb	100.39716
Query.BL00050001.pdb	145.29884
Query.BL00060001.pdb	91.78841
Query.BL00070001.pdb	862.19958
Query.BL00080001.pdb	4366.67725
Query.BL00090001.pdb	97.62597
Query.BL00100001.pdb	2683.79907

**Aminoácidos: 144 – 157**

Filename	molpdf
Query.BL00010001.pdb	128.38635
Query.BL00020001.pdb	221.58383
Query.BL00030001.pdb	130.48824
Query.BL00040001.pdb	3342.73364
Query.BL00050001.pdb	106.47005
Query.BL00060001.pdb	68.07918
Query.BL00070001.pdb	92.59080
Query.BL00080001.pdb	9662.83398
Query.BL00090001.pdb	749.05597
Query.BL00100001.pdb	11642.42969

**Aminoácidos: 163 – 173**

Filename	molpdf
Query.BL00010001.pdb	2389.58325
Query.BL00020001.pdb	1935.60010
Query.BL00030001.pdb	1927.59753
Query.BL00040001.pdb	3290.66406
Query.BL00050001.pdb	3847.09155
Query.BL00060001.pdb	2550.73364
Query.BL00070001.pdb	1926.00012
Query.BL00080001.pdb	2512.98438
Query.BL00090001.pdb	4345.14502
Query.BL00100001.pdb	2701.79199

**Aminoácidos: 196 – 206**

Filename	molpdf
Query.BL00010001.pdb	1834.75903
Query.BL00020001.pdb	303.47089
Query.BL00030001.pdb	2758.85669
Query.BL00040001.pdb	2517.32007
Query.BL00050001.pdb	2500.61230
Query.BL00060001.pdb	4248.29883
Query.BL00070001.pdb	4217.97998
Query.BL00080001.pdb	211.59698
Query.BL00090001.pdb	281.76852
Query.BL00100001.pdb	1217.17737

**Aminoácidos: 293 – 303**

Filename	molpdf
Query.BL00010001.pdb	232.60037
Query.BL00020001.pdb	117.70888
Query.BL00030001.pdb	3063.49463
Query.BL00040001.pdb	1012.01965
Query.BL00050001.pdb	820.83154
Query.BL00060001.pdb	204.92577
Query.BL00070001.pdb	3941.67944
Query.BL00080001.pdb	761.40912
Query.BL00090001.pdb	3398.45142
Query.BL00100001.pdb	86.94650

**Aminoácidos: 326 – 331**

Filename	molpdf
Query.BL00010001.pdb	36.67420
Query.BL00020001.pdb	1046.28027
Query.BL00030001.pdb	59.31277
Query.BL00040001.pdb	29.10810
Query.BL00050001.pdb	68.56017
Query.BL00060001.pdb	9.08036
Query.BL00070001.pdb	25.37222
Query.BL00080001.pdb	58.60068
Query.BL00090001.pdb	52.19076
Query.BL00100001.pdb	63.19436

**Aminoácidos: 415 – 430**

Filename	molpdf
Query.BL00010001.pdb	74.29930
Query.BL00020001.pdb	43.23658
Query.BL00030001.pdb	88.02527
Query.BL00040001.pdb	9464.72754
Query.BL00050001.pdb	98.38969
Query.BL00060001.pdb	92.65536
Query.BL00070001.pdb	70.81356
Query.BL00080001.pdb	93.42401
Query.BL00090001.pdb	90.57508
Query.BL00100001.pdb	86.31180

**Aminoácidos: 1 – 9**

Filename	molpdf
Query.BL00010001.pdb	54.55104
Query.BL00020001.pdb	47.76587
Query.BL00030001.pdb	2691.39673
Query.BL00040001.pdb	1882.87292
Query.BL00050001.pdb	52.07455
Query.BL00060001.pdb	516.66888
Query.BL00070001.pdb	28.60068
Query.BL00080001.pdb	51.54051
Query.BL00090001.pdb	55.02054
Query.BL00100001.pdb	2568.91724

### 8.3 IDENTIFICACIÓN DE LA TRIADA CATALÍTICA DE LA LIPASA 2-17

#### 8.3.1 Secuencia lipasa 2-17 nativa

Lipasa Nativa:

MPILPVPSLNALLTKLTKVKKTAAKNVHQHHILHHTLKGLDNLPTLLERVNHQLKASTAEQYPLADAHRLRILAVSNKL  
KRPLALDKLPNLRQKFGTDAVSLQAPSVWQQNAEASGENSENAVSWQDKTIANADGGDMTVRCYQQSKQNNNEGKST  
DEVAMLFFHGGGFCIGDIDTHHEFCHTVCAQTGWAVVSVDYRMAPEYPAPTALKDCLSAYAWLTEHSQSLGALPSRIV  
LSGDSAGGCLAALVAQQVIKPIDALWQDNNQATETDKKANDTFKSLADLPRPLAQLPLYPVTDYEAEPYPSWELYGEG  
LLDHNDANVFNTAYTQHSGLMQSHPLISVMHGDNAHLCPYSVVVAELDILRDEGLAYAELLQKEGVKVETYTVLGAPH  
GFINLMSVHQGLGNQNTNYIISFACFVQNLLETSEGDEPNL

#### 8.3.2 Secuencias mutaciones lipasa 2-17

Mutación S239N:

MPILPVPSLNALLTKLTKVKKTAAKNVHQHHILHHTLKGLDNLPTLLERVNHQLKASTAEQYPLADAHRLRILAVSNKL  
KRPLALDKLPNLRQKFGTDAVSLQAPSVWQQNAEASGENSENAVSWQDKTIANADGGDMTVRCYQQSKQNNNEGKST  
DEVAMLFFHGGGFCIGDIDTHHEFCHTVCAQTGWAVVSVDYRMAPEYPAPTALKDCLSAYAWLTEHSQSLGALPSRIV  
LSGDNAGGCLAALVAQQVIKPIDALWQDNNQATETDKKANDTFKSLADLPRPLAQLPLYPVTDYEAEPYPSWELYGEG  
LLDHNDANVFNTAYTQHSGLMQSHPLISVMHGDNAHLCPYSVVVAELDILRDEGLAYAELLQKEGVKVETYTVLGAPH  
GFINLMSVHQGLGNQNTNYIISFACFVQNLLETSEGDEPNL

Mutación D361A:

MPILPVPSLNALLTKLTKVKKTAAKNVHQHHILHHTLKGLDNLPTLLERVNHQLKASTAEQYPLADAHRLRILAVSNKL  
KRPLALDKLPNLRQKFGTDAVSLQAPSVWQQNAEASGENSENAVSWQDKTIANADGGDMTVRCYQQSKQNNNEGKST  
DEVAMLFFHGGGFCIGDIDTHHEFCHTVCAQTGWAVVSVDYRMAPEYPAPTALKDCLSAYAWLTEHSQSLGALPSRIV  
LSGDSAGGCLAALVAQQVIKPIDALWQDNNQATETDKKANDTFKSLADLPRPLAQLPLYPVTDYEAEPYPSWELYGEG  
LLDHNDANVFNTAYTQHSGLMQSHPLISVMHGDNAHLCPYSVVVAELAILRDEGLAYAELLQKEGVKVETYTVLGAPH  
GFINLMSVHQGLGNQNTNYIISFACFVQNLLETSEGDEPNL

Mutación D361N:

MPILPVPSLNALLTKLTKVKKTAAKNVHQHHILHHTLKGLDNLPTLLERVNHQLKASTAEQYPLADAHRLRILAVSNKL  
KRPLALDKLPNLRQKFGTDAVSLQAPSVWQQNAEASGENSENAVSWQDKTIANADGGDMTVRCYQQSKQNNNEGKST  
DEVAMLFFHGGGFCIGDIDTHHEFCHTVCAQTGWAVVSVDYRMAPEYPAPTALKDCLSAYAWLTEHSQSLGALPSRIV  
LSGDSAGGCLAALVAQQVIKPIDALWQDNNQATETDKKANDTFKSLADLPRPLAQLPLYPVTDYEAEPYPSWELYGEG  
LLDHNDANVFNTAYTQHSGLMQSHPLISVMHGDNAHLCPYSVVVAELNILRDEGLAYAELLQKEGVKVETYTVLGAPH  
GFINLMSVHQGLGNQNTNYIISFACFVQNLLETSEGDEPNL

Mutación D361E:

MPILPVPSLNALLTKLTKVKKTAAKNVHQHHILHHTLKGLDNLPTLLERVNHQLKASTAEQYPLADAHRLILAVSNKL  
 KRPLALDKLPNLRQKFGTDAVSLQAPSVWQQNAEASGNSENAVSWQDKTIANADGGDMTVRCYQQSKQNNNEGKST  
 DEVAMLFFHGGGFCIGDIDTHHEFCHTVCAQTGWAVVSVDYRMAPEYPAPTALKDCLSAYAWLTEHSQSLGALPSRIV  
 LSGDSAGGCLAALVAQQVIKPIDALWQDNNQATETDKKANDTFKKSADLPRPLAQLPLYPVTDYEAEPYPSWELYGEGEGL  
 LLDHNDANVFNTAYTQHSGLMQSHPLISVMHGDNAHLCPYSVVVAELRDEGLAYAELLQKEGVKVETVTVLGAPH  
 GFINLMSVHQGLGNQNTNYIIKSFACFVQNLTTSEGDEPNL

Mutación H391N

MPILPVPSLNALLTKLTKVKKTAAKNVHQHHILHHTLKGLDNLPTLLERVNHQLKASTAEQYPLADAHRLILAVSNKL  
 KRPLALDKLPNLRQKFGTDAVSLQAPSVWQQNAEASGNSENAVSWQDKTIANADGGDMTVRCYQQSKQNNNEGKST  
 DEVAMLFFHGGGFCIGDIDTHHEFCHTVCAQTGWAVVSVDYRMAPEYPAPTALKDCLSAYAWLTEHSQSLGALPSRIV  
 LSGDSAGGCLAALVAQQVIKPIDALWQDNNQATETDKKANDTFKKSADLPRPLAQLPLYPVTDYEAEPYPSWELYGEGEGL  
 LLDHNDANVFNTAYTQHSGLMQSHPLISVMHGDNAHLCPYSVVVAELDILRDEGLAYAELLQKEGVKVETVTVLGAPN  
 GFINLMSVHQGLGNQNTNYIIKSFACFVQNLTTSEGDEPNL

### 8.3.3 Energías de Afinidad Docking Molecular

Tabla IX: Energía de activación *Docking* Molecular Lipasa nativa y diferentes sustratos

Sustrato	Energía de Afinidad [Kcal/mol]	ΔEnergía de Afinidad [Kcal/mol]
Ác. Caproico	-6,9	-
Ác. Caprílico	-4,7	2,2
Ác. Cáprico	-3,8	3,1
Ác. Palmítico	-2,3	4,6

Tabla X: Energía de activación *Docking* Molecular Lipasa nativa y variantes para sustrato de 6vcarbonos.

Docking Molecular de Variantes de la Lipasa 2-17 y sustrato Ác. Caproico		
Variante	Energía de Afinidad [Kcal/mol]	ΔEnergía de Afinidad [Kcal/mol]
Lip Nativa	-6,9	-
S239N	-2,1	4,8
D361A	-1,9	5
D361E	-2,6	4,3
D361N	-1,1	5,8
H391N	-1,3	5,6

Tabla XI: Energía de activación *Docking* Molecular Lipasa nativa y variantes para sustrato de 8 carbonos.

Docking Molecular de Variantes de la Lipasa 2-17 y sustrato Ác. Caprílico		
Variante	Energía de Afinidad [Kcal/mol]	$\Delta$ Energía de Afinidad [Kcal/mol]
Lip Nativa	-4,7	-
S239N	-1,1	3,6
D361A	-0,8	3,9
D361E	-1,5	3,2
D361N	-2	2,7
H391N	-0,9	3,8

Tabla XII: Energía de activación *Docking* Molecular Lipasa nativa y variantes para sustrato de 10 carbonos.

Docking Molecular de Variantes de la Lipasa 2-17 y sustrato Ác. Cáprico		
Variante	Energía de Afinidad [Kcal/mol]	$\Delta$ Energía de Afinidad [Kcal/mol]
Lip Nativa	-3,8	-
S239N	-1,2	2,6
D361A	-0,6	3,2
D361E	-2,2	1,6
D361N	-1,3	2,5
H391N	-0,8	3

Tabla XIII: Energía de activación *Docking* Molecular Lipasa nativa y variantes para sustrato de 10 carbonos.

Docking Molecular de Variantes de la Lipasa 2-17 y sustrato Ác. Palmítico		
Variante	Energía de Afinidad [Kcal/mol]	$\Delta$ Energía de Afinidad [Kcal/mol]
Lip Nativa	-2,3	-
S239N	-0,6	1,7
D361A	-0,4	1,9
D361E	-1,7	0,6
D361N	-0,9	1,4
H391N	-0,7	1,6



## 8.4 GENERACIÓN DE MUTANTES MEDIANTE MUTAGÉNESIS SITIO-DIRIGIDA

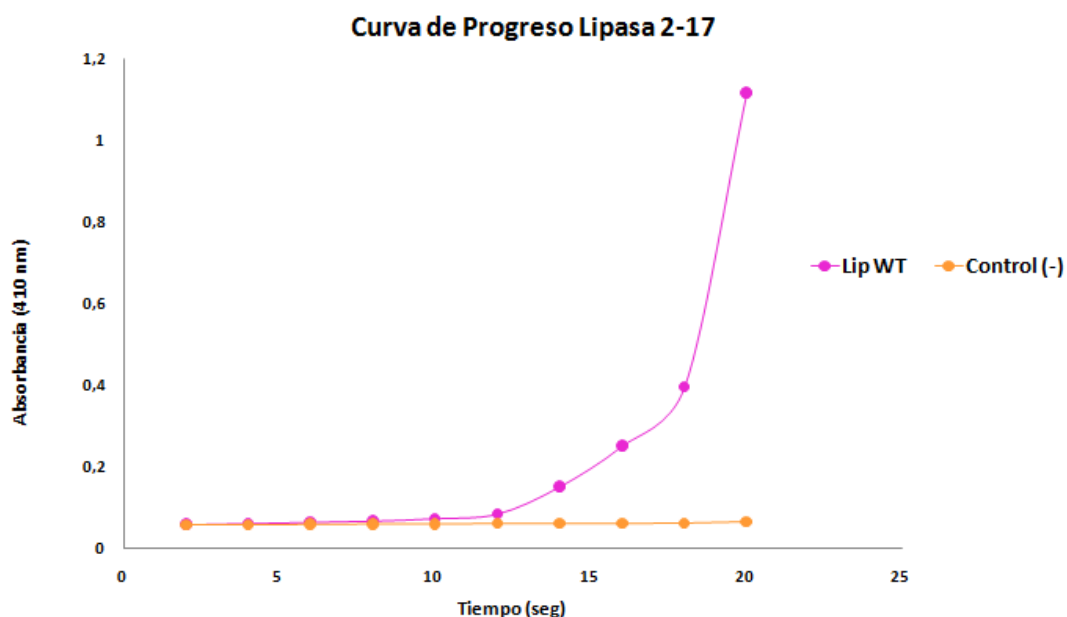
### 8.4.1 Secuencia lipasa M37

>2ORY

```
MSYTKEQLMLAFSYMSYYGITHTGSAKKNAELILKKMKEALKTWKPFQEDDWEVWVWGPVYTMPFTIFNDAMMYVI  
QKKGAEGEYVIAIRGTNPVSDWLFNDFMVSAMKKWPYASVEGRILKISESTSYGLKTLQKLPKSHIPGENKTILQFLN  
EKIGPEGKAKICVTGHSKGGALSSTLALWLKDIQGVKLSQNIDISTIPFAGPTAGNADFADYFDDCLGDQCTRIANSLDIVP  
YAWNTNSLKKLSIYISEQASVKPLLYQRALIRAMIAETKGKYYKQIKAETPPLEGNINPILIEYLVQAAYQHVVGYPELMG  
MMDDIPLTDIFEDAIAGLLHHHHHH
```

## 8.5 ESTUDIO DEL FENÓMENO DE ACTIVACIÓN INTERFACIAL DE LA LIPASA 2-17

### 8.5.1 Curva de Progreso Lipasa 2-17



**Figura 70: Curva de Progreso Lipasa 2-17 para el fenómeno de activación interfacial.** Lipasa Nativa indicada por curva rosa. Control negativo indicado por curva anaranjada.