



# UNIVERSIDAD AUTÓNOMA METROPOLITANA

---

---

DIVISIÓN DE CIENCIAS BÁSICAS E INGENIERÍA  
DEPARTAMENTO DE QUÍMICA  
ÁREA DE BIOFISICOQUÍMICA

## **Efecto del pH sobre la afinidad de quimopapaína con el inhibidor cistatina de pollo: Una asociación enzima-inhibidor modulada por interacción electrostática.**

Tesis que presenta:

**M. en C. Francisco Reyes Espinosa.**

Para obtener el grado académico de

**Doctor en Ciencias (Química).**

Posgrado en Ciencias (Química).

Doctorado Tradicional

Asesor: Dr. José Alfonso Arroyo Reyna

Jurado calificador:

Presidente: Dr. Andrés Hernández Arana

Secretario: Dr. Rafael Arturo Zubillaga Luna

Vocal: Dra. Georgina Regina Garza Ramos Martínez

Vocal: Dr. Horacio Reyes Vivas

Vocal: Dr. Gerardo Pérez Hernández

México, D.F. Junio 2015.



Casa abierta al tiempo

UNIVERSIDAD AUTÓNOMA METROPOLITANA

Unidad Iztapalapa

Fecha : 29/06/2015

Página : 1/1

**CONSTANCIA DE PRESENTACION DE EXAMEN DE GRADO**

La Universidad Autónoma Metropolitana extiende la presente CONSTANCIA DE PRESENTACION DE DISERTACIÓN PÚBLICA de DOCTOR EN CIENCIAS (QUIMICA) del alumno FRANCISCO REYES ESPINOSA, matrícula 206380701, quien cumplió con los 328 créditos correspondientes a las unidades de enseñanza aprendizaje del plan de estudio. Con fecha treinta de junio del 2015 presentó la DEFENSA de su DISERTACIÓN PÚBLICA cuya denominación es:

EFFECTO DEL pH SOBRE LA AFINIDAD DE QUIMOPAPAINA CON EL INHIBIDOR CISTATINA DE POLLO: UNA ASOCIACION ENZIMA-INHIBIDOR MODULADA POR INTERACCION ELECTROSTATICA.

Cabe mencionar que la aprobación tiene un valor de 180 créditos y el programa consta de 483 créditos.

El jurado del examen ha tenido a bien otorgarle la calificación de:

Aprobar

**JURADO**

**Presidente**

**Secretario**

DR. ANDRES HERNANDEZ ARANA

DR. RAFAEL ARTURO ZUBILLAGA LUNA

**Vocal**

**Vocal**

CANCELADO

DR. GERARDO PEREZ HERNANDEZ

DR. HORACIO REYES VIVAS

**Vocal**

  
DRA. GEORGINA REGINA GARZA RAMOS  
MARTINEZ

## RESUMEN DE CONTENIDO.

I. LISTA DE TABLAS.	ii
II. LISTA DE FIGURAS.	iii
III. ABREVIATURAS.	v
IV. ECUACIONES.	vi
V. ÍNDICE DE CONTENIDO.	vi
VI. RESUMEN.	x
VII. AGRADECIMIENTOS.	xii

### I. LISTA DE TABLAS.

<b>Tabla 1.1</b> Clasificación de enzimas.	1
<b>Tabla 1.2</b> Clasificación de proteasas.	2
<b>Tabla 1.3</b> Propiedades de proteinasas contenidas en látex de Carica papaya L.	5
<b>Tabla 1.4</b> Propiedades fisicoquímicas de proteínas presentes en clara de huevo.	8
<b>Tabla 1.5</b> Constantes de inhibición ( $K_i$ ) y disociación ( $K_d$ ) de sistemas cisteinproteasa–cistatina.	13
<b>Tabla 4.1</b> Razón de la medición de absorbancia en UV de proteínas ( $Abs^{280/260}$ y $Abs^{280/292}$ ).	42
<b>Tabla 4.2</b> Estructura secundaria de SCM-quimopapaína y CEW a diferentes valores de pH (análisis efectuado con el método CONTIN).	44
<b>Tabla 4.3</b> Evidencia experimental de la formación del complejo SCM-quimopapaína–CEW a los diferentes valores de pH en estudio.	46
<b>Tabla 4.4</b> Constante de equilibrio para la unión de SCM-quimopapaína y CEW determinada a diferentes valores de pH a 25 °C.	48
<b>Tabla 4.5</b> Interacciones relevantes localizadas en la interfase del modelo quimopapaína–CEW.	49
<b>Tabla 4.6</b> Propiedades estructurales y fisicoquímicas de residuos ionizables identificados en la interfase del heterodímero quimopapaína–CEW.	51
<b>Tabla 4.7</b> Parámetros obtenidos del mejor ajuste de los datos experimentales de $K_b$ en función del pH al modelo <i>MPLM</i> .	54
<b>Tabla 4.8</b> Residuos de aminoácidos identificados en el subsitio $S_2$ de cisteinproteasas.	56

<b>Tabla 4.9</b> Residuos ionizables y propiedades estructurales de la interfase de heterodímeros homólogos al complejo cisteinproteasa–cistatinas.	58
<b>Tabla 4.10</b> Residuos <i>hot spots</i> identificados en los complejo homólogos cisteinproteasa–cistatina.	59
<b>Tabla 5.1</b> Características generales de la interacción de estructuras homólogas cisteinproteasa–cistatina.	70
<b>Tabla A.1</b> Métodos implementados para la purificación de CEW.	80
<b>Tabla E.1</b> Alineamiento de secuencia primaria de estructuras homólogas cisteinproteasa–cistatina e identificación de residuos en la interfase.	96
<b>Tabla E.2</b> Composición de residuos de aminoácidos de proteínas que componen las estructuras homólogas cisteinproteasa–cistatina en comparación.	98
<b>Tabla E.3</b> Propiedades estructurales de las proteínas que componen los heterodímeros indicados en Tabla E.1.	99
<b>Tabla F.1</b> Análisis computacional <i>alanine scanning</i> de complejos homólogos cisteinproteasa–cistatina.	101

## II. LISTA DE FIGURAS.

<b>Figura 1.1</b> Modelo de Schechter y Berger (1967).	3
<b>Figura 1.2</b> Mecanismo catalítico de CPs (Ejemplificado con papaína).	4
<b>Figura 1.3</b> Estructura de quimopapaína (PDB 1YAL dibujada con el programa YASARA).	6
<b>Figura 1.4</b> Clasificación de cistatinas.	7
<b>Figura 1.5</b> Estructura de cistatina de pollo (PDB 1A67 dibujada con el programa YASARA).	9
<b>Figura 3.1</b> Carboximetilación (SCM-) de cisteína con ácido iodo acético.	15
<b>Figura 3.2</b> Procedimiento de purificación de SCM-quimopapaína.	15
<b>Figura 3.3</b> Tratamiento de clara de huevo.	19
<b>Figura 3.4</b> Preparación de SCM-papaína (solución acoplante).	20
<b>Figura 3.5</b> Cromatografía de afinidad para la purificación de CEW (Cromatografía en <i>batch</i> ).	22
<b>Figura 3.6</b> Representación de la hidrólisis del sustrato BApNA por papaína.	23
<b>Figura 3.7</b> Diagrama termodinámico que ilustra el procedimiento estándar para calcular la contribución electrostática a la energía libre de unión de dos moléculas con cuerpo rígido.	29

<b>Figura 4.1</b> Cromatograma de extracto de quimopapaína.	32
<b>Figura 4.2</b> Cromatograma de SCM-quimopapaína.	33
<b>Figura 4.3</b> Espectro de masas MALDI-TOF de la SCM-quimopapaína.	33
<b>Figura 4.4</b> Propiedades espectrales de SCM-quimopapaína	34
<b>Figura 4.5</b> Método de purificación de cistatina a partir de clara de huevo de pollo.	35
<b>Figura 4.6</b> Cromatograma de solución de cistatina obtenida de cromatografía de afinidad.	36
<b>Figura 4.7</b> Cromatograma de exclusión molecular CEW.	37
<b>Figura 4.8</b> Cinéticas típicas de la hidrólisis de sustrato BApNA por papaína.	38
<b>Figura 4.9</b> Gel de electroforesis SDS-PAGE de CEW.	39
<b>Figura 4.10</b> Espectro de masas MALDI-TOF de CEW.	39
<b>Figura 4.11</b> Propiedades espectrales de CEW.	40
<b>Figura 4.12</b> Efecto del pH sobre propiedades espectrales de SCM-quimopapaína y CEW.	41
<b>Figura 4.13</b> Espectros de CD de SCM-quimopapaína en la región UV lejano (190-250 nm) a diferentes valores de pH.	43
<b>Figura 4.14</b> Espectros de CD de CEW en la región UV lejano (190-250 nm) a diferentes valores de pH.	43
<b>Figura 4.15</b> Fluorescencia intrínseca de SCM-quimopapaína y CEW.	45
<b>Figura 4.16</b> Titulación de SCM-quimopapaína con CEW monitoreado por mediciones de fluorescencia intrínseca de proteínas.	47
<b>Figura 4.17</b> Evolución de la desviación del valor cuadrático medio (RMSD) estimado para los C $\alpha$ durante la simulación MD en solvente explícito en un conjunto TNP a 300 K y 1 bar. La estructura de referencia para el cálculo fue la estructura inicial del heterodímero quimopapaína–CEW.	49
<b>Figura 4.18</b> Modelo estructural del complejo quimopapaína–CEW.	50
<b>Figura 4.19</b> Efecto del pH sobre la constante de unión ( $K_b$ ) de quimopapaína con CEW.	52
<b>Figura 4.20</b> Comparación de la energía libre de unión experimental ( $\Delta G_b$ ) y teórico ( $\Delta G_{b,elec}$ ) del complejo enzima–inhibidor con respecto a la variación de pH.	55
<b>Figura 5.1</b> Dependencia del número de iones H $^+$ ( $q$ ) transferidos hacia la solución amortiguadora de pH con respecto a la variación de pH.	67
<b>Figura B.1</b> Cromatograma de CEW purificada.	82

### III. ABREVIATURAS.

Abs	Absorbancia.
ASP	Área de superficie de contacto.
CAS	Computational Alanine Scanning.
CEW	Cistatina de clara de huevo de pollo.
CPs	Cisteinproteasas.
Da	Daltons.
DC	Dicroísmo circular.
MD	Dinámica Molecular.
$F_P$	Fluorescencia de SCM- quimopapaína libre
$F_C$	Fluorescencia de CEW libre.
$F_M$	Fluorescencia de la mezcla de SCM-quimopapaína y CEW
HSP	Homología de secuencia primaria.
$I$	Fuerza iónica.
$K_b$	Constante de unión.
$K_d$	Constante de disociación.
$K_i$	Constante de inhibición.
$M$	Número de grupos isoacídicos.
$M$	Número de residuos que constituyen el grupo isoacídico.
M	Masa molar.
NMR	Resonancia Magnética Nuclear.
<i>MPLM</i>	Modelo modificado de transferencia de protón vinculado a la asociación.
$N_{HB}$	Número de puentes de hidrógeno a través de la interfase.

$N_{IP}$	Número de pares iónicos a través de la interfase.
PDB	Protein Data Bank.
pH	Logaritmo negativo de la concentración de protones.
pI	Punto isoelectrico.
$pK_a$	Logaritmo negativo de la constante de disociación ácida.
$q$	Número de iones $H^+$ .
RMSD	Desviación del valor cuadrático medio.
SAS	Área de superficie accesible al solvente.
SCM	S-carboximetil.
SDS-PAGE	Electroforesis en gel de poliacrilamida en presencia de dodecil- sulfato de sodio.
UV	Ultravioleta.
$\Delta G_b$	Energía libre de unión.
$\Delta G_{b,elec}$	Energía electrostática de unión.
$\Delta G_{coul}$	Cambio en la energía coulombica.
$\Delta G_{solv}$	Cambio en la energía de solvatación.
$[\theta]_{MRW}$	Elipticidad molar por residuo medio.
$\Delta\Delta G_{bind}^{CAS}$	Cambio en la energía libre de unión, en el análisis CAS
$\Delta G_{bind}^{CAS}$	Cambio en la estabilidad de la proteína sola, en el análisis CAS.
$\lambda$	Longitud de onda
Y	Rendimiento.

#### IV. ECUACIONES.

3.1	Elipticidad molar por residuo medio, $[\theta]_{MRW}$ .	18
3.2	Modelo matemático de unión bimolecular para determinar $K_d$ .	26
3.3	Ecuación del modelo <i>MPLM</i> (Modelo modificado de transferencia de protón vinculado a la asociación).	31
3.4	Dependencia del número de iones $H^+$ ( $q$ ) transferidos a partir del sistema de estudio hacia el amortiguador de pH.	31
4.1	Ecuación del modelo <i>MPLM</i> , para tres grupos isoacídicos.	53

#### V. ÍNDICE DE CONTENIDO.

1.	INTRODUCCIÓN.	1
1.1.	Generalidades de la clasificación de las enzimas.	1
1.2.	Cisteinproteasas.	2
1.2.1.	Mecanismo catalítico de cisteinproteasas.	3
1.3.	Proteinasas obtenidas a partir del látex de <i>Carica papaya L.</i>	5
1.4.	Cistatinas: Inhibidores de cisteinproteasas.	7
1.5.	Cistatina de clara de huevo de pollo.	8
1.6.	Interacción proteína–proteína y su dependencia con el pH.	10
1.7.	Importancia y aplicaciones del estudio de cisteinproteasas y sus inhibidores.	12
2.	OBJETIVOS.	14
2.1.	Objetivos generales.	14
2.2.	Objetivos particulares.	14
3.	MATERIALES Y MÉTODOS.	15
3.1.	Método de purificación de SCM-quimopapaína.	15
3.1.1.	Pruebas de pureza de SCM-quimopapaína.	16
	A. <i>Recromatografía de intercambio catiónico.</i>	16
	B. <i>Espectrometría de masas.</i>	16

3.1.2. Propiedades espectrales de quimopapaína purificada: Espectros de fluorescencia y dicroísmo circular (DC).	17
A. <i>Espectros de fluorescencia.</i>	17
B. <i>Espectros de DC.</i>	18
3.2. Método de purificación de cistatina de pollo (CEW).	18
3.2.1. Preparación de clara de huevo.	19
3.2.2. Inmovilización de SCM-papaína sobre sefarosa 4B para cromatografía de afinidad.	20
3.2.3. Cromatografía de afinidad para la purificación de CEW.	21
3.2.4. Cromatografía de exclusión molecular para la purificación de CEW.	22
3.2.5. Ensayos de actividad inhibitoria y pruebas de pureza de CEW.	23
A. <i>Ensayo de actividad inhibitoria de CEW.</i>	23
B. <i>Electroforesis en gel de poliacrilamida en presencia de dodecil - sulfato de sodio (SDS-PAGE).</i>	24
C. <i>Espectrometría de masas.</i>	24
3.2.6. Propiedades espectrales de CEW purificada: Espectros de fluorescencia y dicroísmo circular (DC).	24
A. <i>Espectros de fluorescencia.</i>	24
B. <i>Espectros de DC.</i>	24
3.3. Métodos espectroscópicos.	25
3.3.1. Espectros de absorbancia en la región UV-Visible y dicroísmo circular (DC) en la región UV-lejano en el intervalo de pH de 3.5 a 10.0, a 25 °C.	25
3.3.2. Determinación de la constante de disociación ( $K_d$ ) por medio del método de titulación fluorimétrica del sistema SCM-quimopapaína-CEW.	25
3.4. Métodos computacionales.	27
3.4.1. Modelado de la interacción quimopapaína-CEW y simulaciones con Dinámica Molecular (MD).	27
3.4.2. Cálculo de la energía electrostática de unión ( $\Delta G_{b,elec}$ ) para el complejo quimopapaína-CEW.	28
3.4.3. Comparación del sistema quimopapaína-CEW con estructuras homólogas cisteinproteasa-cistatina.	29
3.5. Modelo modificado de transferencia de protón vinculado a la asociación: <i>MPLM</i> (Crnogorac <i>et al.</i> , 2001).	30



4. RESULTADOS.	32
4.1. Purificación de SCM-quimopapaína.	32
4.1.1. Pruebas de pureza de SCM-quimopapaína.	33
4.1.2. Propiedades espectrales de SCM-quimopapaína purificada.	34
4.2. Purificación de CEW.	35
4.2.1. Ensayos de actividad inhibitoria y pruebas de pureza de CEW.	37
4.2.2. Propiedades espectrales de CEW purificada.	40
4.3. Cambios espectroscópicos que acompañan la interacción de SCM-quimopapaína y CEW a diferentes valores de pH, a 25 °C.	41
4.3.1. Cambios espectrales en absorbancia en la región UV-Visible y DC en la región UV-lejano a pH de 3.5 a 10.0.	41
4.3.2. Cambios espectroscópicos que acompañan la interacción de SCM-quimopapaína y CEW: Constante unión ( $K_b$ ) como función de pH.	45
4.4. Modelo molecular de interacción quimopapaína–CEW.	49
4.5. Dependencia de $K_b$ con respecto a la variación de pH y modelo <i>MPLM</i> (Crnogorac <i>et al.</i> , 2001).	52
4.6. Comparación de energía libre de unión ( $\Delta G_b$ ) y energía electrostática de unión ( $\Delta G_{b,elec}$ ).	55
4.7. Comparación estructural de la interacción quimopapaína–CEW con otras estructuras homólogas cisteinproteasas–cistatinas	56
A. <i>Composición del subsitio <math>S_2</math> de cisteinproteasas comparadas.</i>	56
B. <i>Aspectos estructurales de la interacción cisteinproteasas–cistatinas.</i>	57
C. <i>Comparación de interfases de complejos homólogos cisteinproteasa–cistatina.</i>	57
D. <i>Análisis computacional alanine scanning de complejos homólogos cisteinproteasa–cistatina: Residuos hot spots.</i>	58
5. DISCUSIÓN.	60
5.1. Purificación de SCM-quimopapaína.	60
5.2. Purificación de CEW.	60
A. <i>Método de purificación de cistatina de pollo a partir de clara de huevo.</i>	60
a) <i>Tratamiento de clara de huevo.</i>	61
b) <i>Cromatografía de afinidad.</i>	61
c) <i>Cromatografía por exclusión molecular.</i>	62

B. <i>Ensayos de actividad inhibitoria.</i>	62
C. <i>Propiedades espectrales y pruebas de pureza de CEW purificada.</i>	63
5.3. Propiedades espectrales de SCM-quimopapaína y CEW: Espectros de absorbancia en la región UV-Visible, fluorescencia y DC en la región UV-lejano a valores de pH en estudio.	64
A. <i>Espectros de absorbancia en la región UV-Visible.</i>	64
B. <i>Espectros de fluorescencia.</i>	65
C. <i>Espectros de DC en la región UV-lejano.</i>	65
5.4. Determinación de la constante de unión ( $K_b$ ).	65
5.5. Dependencia de $K_b$ con respecto a la variación de pH.	66
5.6. Comparación estructural de la interacción quimopapaína–CEW con otras estructuras homólogas cisteinproteasa–cistatina.	68
A. <i>Composición del subsitio <math>S_2</math> de cisteinproteasas comparadas.</i>	68
B. <i>Aspectos estructurales de la inhibición de cisteinproteasas por cistatinas.</i>	68
C. <i>Análisis computacional alanine scanning de complejos homólogos cisteinproteasa–cistatina: Residuos hot spots.</i>	69
D. Características de la interacción de estructuras homólogas cisteinproteasa–cistatina.	69
6. CONCLUSIONES	71
7. PERSPECTIVAS.	72
8. REFERENCIAS.	73
9. APÉNDICES	80
A. Métodos desarrollados para purificar cistatina de pollo.	80
B. Cromatografía de intercambio aniónico de CEW purificada.	82
C. Modelo para la determinación de la constante de disociación ( $K_d$ ) para un sistema proteína-ligando mediante el método de titulación fluorimétrica.	83
D. Derivación del modelo de transferencia de protón vinculado a la unión: <i>MPLM</i> (Crnogorac <i>et al.</i> , 2001) desarrollado para tres grupos isoacídicos.	87
E. Propiedades estructurales quimopapaína, CEW y proteínas de complejos homólogos cisteinproteasa–cistatina.	96
F. Residuos <i>hot spots</i> en la interface de complejos homólogos cisteinproteasa–cistatina: Computational <i>alanine scanning</i> of protein–protein.	100

## VI. RESUMEN.

En el presente trabajo se estudió la dependencia de la constante de unión ( $K_b$ ) de la quimopapaína -una cisteinproteasa representativa tipo papaína- con el inhibidor cistatina de pollo (CEW, denotado en el presente trabajo por las siglas de chicken egg white) con respecto a la variación del pH.

Los estudios efectuados con métodos espectroscópicos se realizaron con quimopapaína teniendo bloqueada su Cys catalítica (carboximetilada) y cistatina de pollo en su forma cortada en el extremo amino terminal en Gly<sup>9</sup>, resultado de su purificación a partir de clara de huevo de pollo. El análisis de los espectros de dicroísmo circular en la región ultravioleta lejano de las proteínas individuales en el intervalo de pH de [3.5–10.0] a 25 °C, no revela evidencia de un cambio significativo en estructura secundaria de proteína. En los espectros de absorbancia en la región ultravioleta cercano a valores de pH alcalinos, no se observó señal de la ionización de la cadena lateral de Tyr. Esta observación se confirmó en los espectros de emisión de fluorescencia de las proteínas individuales y de la mezcla.

La unión de quimopapaína con cistatina de pollo presenta una moderada afinidad, el sistema presenta una  $K_b$  de  $(5.71 \pm 0.27) \times 10^6 \text{ M}^{-1}$  (una  $K_d$  de  $(176 \pm 8) \text{ nM}$ ) a pH 7.4, 25° C ( $K_d$  fue determinada por el método de titulación fluorimétrica y  $K_b$  calculada como:  $K_b = K_d^{-1}$ ).

Mediante el método de titulación fluorimétrica se determinó la constante de disociación ( $K_d$ ) del sistema a 14 diferentes valores de pH a una fuerza iónica ( $I$ ) de 0.11 M en el intervalo [6.0–10.0] y a 5 diferentes valores de pH con una  $I$  de 0.011 M en el intervalo [3.5–6.5]. Estos experimentos se realizaron en presencia de su respectivo amortiguador de pH a 25 °C.

La  $K_d$  determinada a dos distintas fuerzas iónicas ( $I = 0.011 \text{ M}$  y  $I = 0.11 \text{ M}$ ) a valores de pH de 6.0 y 6.5, mostró una diferencia de aproximadamente diez veces mayor en magnitud, evidencia de que la asociación quimopapaína–cistatina de pollo es modulada por interacción electrostática.

La variación de  $K_b$  (calculada como  $K_b = K_d^{-1}$ ) con el pH se logró describir apropiadamente con el modelo modificado de transferencia de protón (*MPLM*), propuesto por Crnogorac *et al.* (2001), con la presencia de tres grupos ionizables isoacídicos.

Con el propósito de elucidar mejor el mecanismo de unión se construyó un modelo de la estructura tridimensional del complejo quimopapaína–CEW. El análisis de la distribución de residuos cargados sobre la interfase del complejo quimopapaína–cistatina a un radio de corte  $\leq 5$  Å de las cadenas, reveló la presencia de ocho residuos ionizables sobre quimopapaína (Y<sup>61</sup>, K<sup>64</sup>, Y<sup>67</sup>, K<sup>139</sup>, K<sup>145</sup>, K<sup>156</sup>, D<sup>158</sup> y H<sup>159</sup>) y cinco sobre cistatina (E<sup>19</sup>, G<sup>9</sup><sub>N-terminal</sub>, R<sup>52</sup>, K<sup>59</sup> y Y<sup>100</sup>).

El cálculo teórico del componente electrostático de la energía libre de unión ( $\Delta G_{b,elec}$ ) del complejo quimopapaína–CEW a los diferentes valores de pH examinados describe cualitativamente el comportamiento experimental observado de la energía libre de unión ( $\Delta G_b$ ) como función del pH.  $\Delta G_{b,elec}$  se calculó empleando un modelo con solvente implícito basado en la resolución de la ecuación de Poisson-Boltzmann.

Durante la asociación de la quimopapaína con cistatina de pollo, las interacciones electrostáticas juegan un papel importante. Estas interacciones son moduladas por el pH del entorno que dicta su estado de protonación de acuerdo con el  $pK_a$  de los residuos ionizables localizados en la interfase del heterodímero y en las proteínas separadas, por lo tanto, también determina la energía de unión electrostática. Proponemos que los residuos que participan en el par iónico entre cadenas, Lys<sup>139</sup> de quimopapaína y Glu<sup>19</sup> de CEW, así como Tyr<sup>61</sup> y Tyr<sup>67</sup> de quimopapaína, son los principales residuos responsables de la dependencia de la afinidad del heterodímero quimopapaína–CEW observada experimentalmente con respecto a la variación del pH.

Finalmente, en base al análisis comparativo de la interfase con estructuras cristalográficas ya reportadas de sistemas homólogos cisteinproteasa–cistatina (PDB: 1STF, 3K9M, 1NB5, 3KSE, 3KFQ), se puede asumir que la dependencia de la constante de unión con respecto a la variación del pH, así como la contribución electrostática a la energía de unión están fuertemente vinculada a la presencia y distribución de grupos ionizables en la interfase, puentes de hidrógeno, pares iónicos y área superficial de contacto.

## **VII. AGRADECIMIENTOS.**

El presente trabajo fue realizado en el Área de Biofísicoquímica del Departamento de Química, División de Ciencias Básicas e Ingeniería, de la Universidad Autónoma Metropolitana, Unidad Iztapalapa, bajo la asesoría del Dr. José Alfonso Arroyo Reyna.

Para la realización de éste trabajo se contó con el apoyo económico de CONACYT a través de la Beca de Doctorado con número 28595, con el número de registro (CVU/Becario): 204984/203393.

También se contó con el apoyo económico de COMECYT a través de la Beca para Tesis de Posgrado con número 13BTD0409.

Agradezco el apoyo económico de CONACYT a través de la Beca de Ayudante de Investigador Nacional Nivel III otorgado por el Dr. Andrés Hernández Arana con número de expediente SNI 629, en el periodo Mayo 2013 - Octubre 2015.

Asimismo agradezco a los Doctores que colaboraron en el artículo, producto del presente trabajo de investigación, al Dr. José Alfonso Arroyo Reyna, Ponciano García Gutiérrez, Dra. Iris N. Serratos Alvares y Dr. Rafael A. Zubillaga Luna.

Gracias a los sinodales por sus observaciones y críticas al manuscrito de esta tesis, los cual contribuyó a mejorar su presentación, al Dr. Andrés Hernández Arana, Dr. Rafael A. Zubillaga Luna, Dra. Georgina Regina Garza Ramos Martínez, Dr. Horacio Reyes Vivas y al Dr. Gerardo Pérez Hernández.

El presente trabajo lo dedico con mucho cariño y amor a mis papas, Marcela Espinosa Ramírez y Ángel Reyes Mendoza. Asimismo a mis hermanos, Celina, Enrique, Ricardo y María Teresa.

De todo corazón...Gracias.

M. en C. Francisco Reyes Espinosa.

Si en la lid el destino te derriba;  
si todo en tu camino es cuesta arriba;  
si tu sonrisa es ansia insatisfecha;  
si hay faena excesiva y vil cosecha;  
si a tu caudal se contraponen diques,  
date una tregua...pero... ¡NO CLAUDIQUES!

Rudyard Kipling

## 1. INTRODUCCIÓN.

### 1.1. Generalidades de la clasificación de las enzimas.

Las enzimas son catalizadores de los sistemas biológicos, sus características más sobresalientes son su poder catalítico y alta especificidad, tanto en la reacción que catalizan como en la selección de sustancias reactantes, denominadas sustratos. La catálisis tiene lugar en un centro específico del enzima llamado centro activo [1].

Para dar consistencia a la clasificación de las enzimas, en 1964 la International Union of Biochemistry and Molecular Biology estableció un comité para la nomenclatura de enzimas (NC-IUBMB). La mayor parte de las enzimas se denominan a partir del nombre de los sustratos y las reacciones que catalizan con el sufijo “asa”. Estas reacciones se dividen en seis clases principales (ver Tabla 1.1), para identificar con precisión a las enzimas se les asigna un código de cuatro dígitos precedido por las letras EC (de Enzyme Commission). Para nuestro propósito de estudio, nos conciernen las peptidasas (subclase EC 3.4) también conocido como proteasas, proteinasas o enzimas proteolíticas [1,2].

**Tabla 1.1** Clasificación de enzimas.

Clase.	Tipo de reacción.
1. Oxidoreductasas	Catalizan oxido-reducciones entre dos sustratos.
2. Transferasas	Transferencia de grupo (distinto del hidrógeno).
3. Hidrolasas	Reacciones de hidrólisis de los enlaces tipo éster, éter, péptido, glucosilo, anhídrido de ácido, C-C, C-halógeno o P-N.
4. Liasas	Adición o separación de grupos para formar dobles enlaces.
5. Isomerasas	Isomerización (transferencia intramolecular de grupo).
6. Ligasas	Unión de dos sustratos a expensas de la hidrólisis de ATP.

Las proteasas (enzimas que catalizan la hidrólisis de un enlace peptídico o proteólisis) son de gran importancia en varias funciones celulares, tienen papeles clave en diversos procesos fisiológicos esenciales. Las proteasas están involucradas en procesos digestivos, activación de proenzimas, liberación de péptidos fisiológicamente activos y procesos de inflamación entre otros. Las proteasas se encuentran extensamente distribuidas en células de plantas y animales así como en bacterias y virus [1].

Dentro de las enzimas que se comercializan a nivel industrial en todo el mundo, las proteasas ocupan la posición más relevante (cerca del 60%). Los campos de aplicación son diversos, incluyen tecnología y ciencia de los alimentos, industria farmacéutica y manufactura de detergentes. El gran desarrollo de la ingeniería genética y tecnología de DNA recombinante han permitido la aplicación industrial de estas enzimas, reduciendo costos y ofreciendo una producción a gran escala. En consecuencia, las enzimas obtenidas a partir de plantas han comenzado a ser remplazadas por enzimas recombinantes. Sin embargo, estas proteasas de plantas son de gran interés para el desarrollo de nuevos productos farmacéuticos y aplicaciones biotecnológicas tanto por su actividad proteolítica sobre una extensa variedad de proteínas como por el intervalo óptimo de temperatura y pH [3].

Para el estudio de las proteasas, en 1992 la NC-IUBMB acordó seis clases de mecanismos de acuerdo a su sitio catalítico (Tabla 1.2). Las proteasas se agrupan en categorías de acuerdo al aminoácido esencial en su sitio activo, el intervalo de pH óptimo de actividad, la similitud en secuencia de aminoácidos y el mecanismo de inhibición [2,4,5].

**Tabla 1.2** Clasificación de proteasas.

Proteasas.	Residuo de aminoácido en el sitio activo.	Intervalo de pH óptimo <sup>1</sup> .
Serinproteasas (EC 3.4.21)	Ser; His.	7 – 9
Cisteinproteasas (EC 3.4.22)	Cys.	4 – 7
Aspartilproteasas (EC 3.4.23)	Asp, Tyr.	Menor a 5
Metaloproteasas (EC 3.4.24)	Ion metálico.	7 – 9
Treoninproteasas (EC 3.4.24)	Thr.	8 – 9
Endoproteasas no clasificadas (EC 3.4.99)	Contienen dos o más sitios catalíticos.	

<sup>1</sup>Oliveira *et al.* (2003) [4].

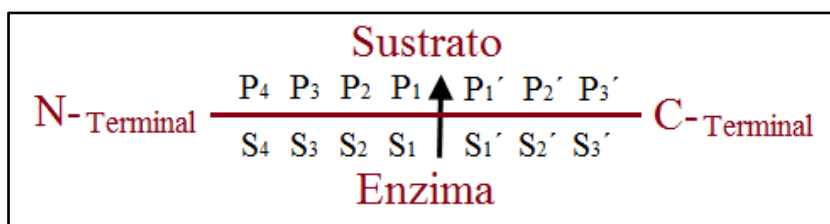
## 1.2. Cisteinproteasas.

Las cisteinproteasas (CPs) son proteínas de masa molecular de alrededor de 21–30 KDa que catalizan la hidrólisis de enlaces peptídico, amida, éster y tioléster. Las CPs están presentes en todos los seres vivos, se han descubierto alrededor de 21 familias, más de la mitad en virus. Muchas de estas enzimas se encuentran en bacterias, hongos, protozoos, plantas (la familia de papaína es la mejor caracterizada) y mamíferos (dos grupos principales: calpains citosólicas y catepsinas lisosomales) [6].



La familia CPs puede ser subdivida en exopeptidasa (por ejemplo: catepsina X, carboxipeptidasa B) y endopeptidasa (por ejemplo: papaína, bromelaina, ficaina, catepsinas). Las endopeptidasas hidrolizan el enlace peptídico interno y las exopeptidasas el enlace próximo al N-terminal o C-terminal del sustrato. El término proteasa es utilizado como sinónimo de péptidasa y puede ser aplicado tanto a exopeptidasas como a endopeptidasas. En tanto proteinasa solo es aplicado a endopeptidasas [5,6].

La especificidad de una peptidasa es descrita por el uso de un modelo conceptual empleado por primera vez en 1967 por Schechter & Berger referente a la papaína [7]. En este modelo, cada subsitio específico es capaz de acoplar en la cadena lateral un residuo de aminoácido del sustrato peptídico. Los subsitios son numerados a partir del sitio catalítico; Para el sustrato:  $P_1, P_2, P_3, P_4, \dots, P_n$  en dirección del N-Terminal y  $P_1', P_2', P_3', \dots, P_n'$  en dirección al C-Terminal y para la enzima: los subsitios son numerados  $S_4, S_3, S_2, S_1, S_1', S_2', S_3'$  respectivamente (Figura 1.1).

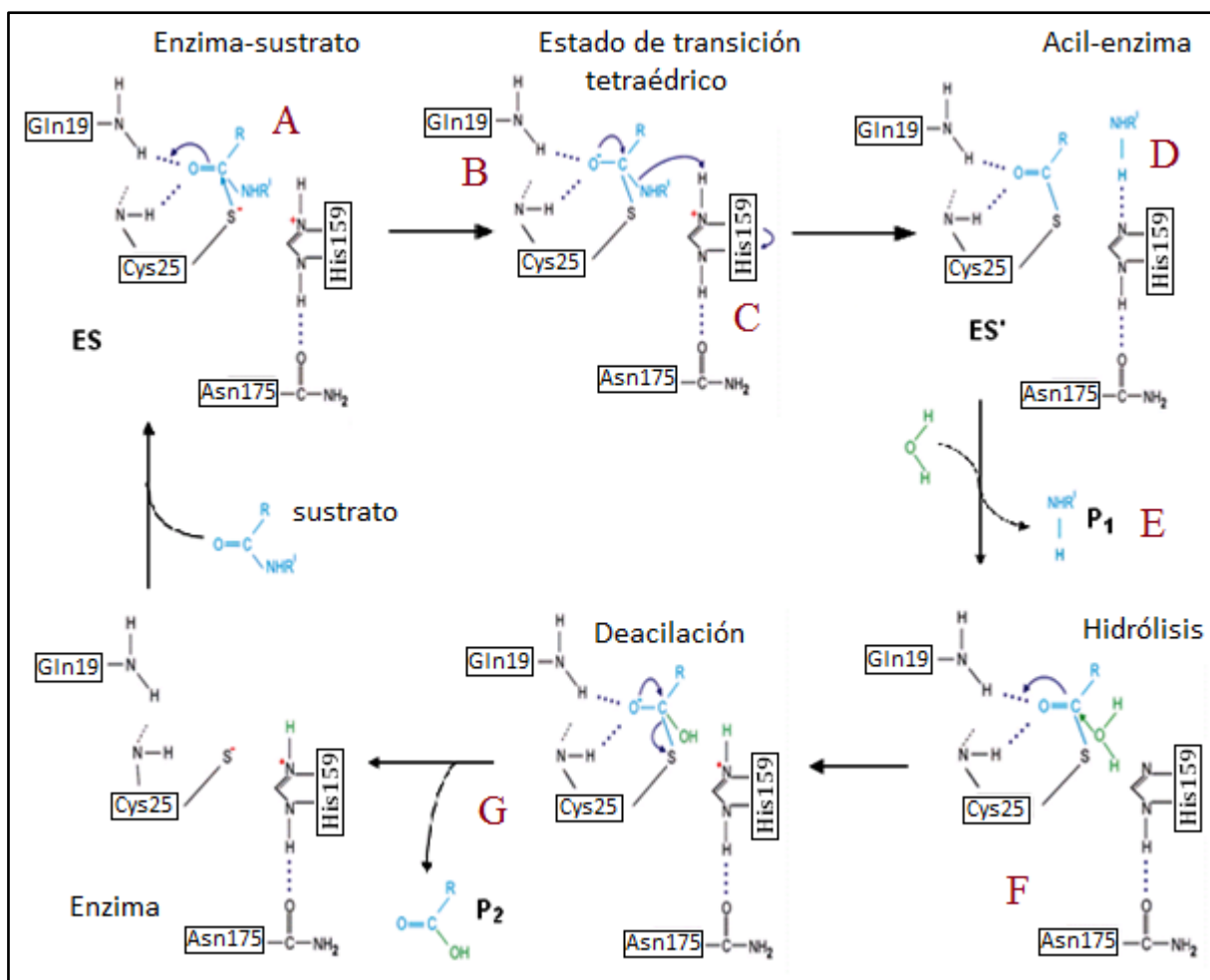


**Figura 1.1** Modelo de Schechter y Berger (1967) [7].

### 1.2.1 Mecanismo catalítico de cisteinproteasas.

El mecanismo general de acción de CPs es ejercida por la diada catalítica formada por los residuos Cys<sup>25</sup> y His<sup>159</sup> (numeración de papaína). El sitio activo es localizado en la hendidura formada por dos dominios, en uno predomina estructura  $\alpha$ -hélice y en otro barril  $\beta$  (por ejemplo; quimopapaína, ver Figura 1.3). La estructura de enzimas pertenecientes a la familia de papaína es muy similar y el mecanismo catalítico es idéntico. Sin embargo, los residuos de aminoácidos adyacentes al sitio catalítico responsables de unir al sustrato pueden diferir y presentar diferencia en su especificidad al sustrato [3].

El modelo mínimo del mecanismo de la hidrólisis de un sustrato amida o éster por CPs involucra la formación de una acil-enzima (tioléster), intermediario ES' y la liberación del producto amina o alcohol ( $P_1$ ), seguida por la hidrólisis (deacilación) de ES' el cual provee el grupo carboxilato ( $P_2$ ) y la regeneración de la enzima (Figura 1.2) [6].



**Figura 1.2** Mecanismo catalítico de CPs (Ejemplificado con papaína). En principio, el ataque del anión tiolato al carbonilo facilita el rompimiento del enlace peptídico y el doble enlace entre el carbono y oxígeno se convierte en enlace sencillo (A). El oxígeno asume una carga neta negativa que permite el primer estado de transición tetraédrico. El oxianión se estabiliza enlazando hidrógenos del grupo NH de la cadena lateral del residuo Gln<sup>19</sup> y Cys<sup>25</sup> de la cadena principal, de lo cual probablemente resulta la formación de la cavidad del oxianión (B). En seguida la rotación de la cadena lateral de His<sup>159</sup>, permite la transferencia de protones desde el catión imidazolío al nitrógeno del enlace peptídico que comienza a ser hidrolizado y ocurre el rompimiento (C). El sustrato recién formado es enlazado a la His<sup>159</sup>, mientras la parte carboxílica del sustrato es enlazado Cys<sup>25</sup> vía enlace tioéster, formando una acil-enzima (D). El siguiente paso de la reacción involucra la disociación de la parte amína del sustrato y su sustitución por una molécula de agua (E). El nitrógeno del imidazol contribuye a la polarización de la molécula de agua que a su vez ataca el carbono del grupo carbonilo de la acil-enzima. Esta es seguida de la formación de un intermediario tetraédrico (F). En la etapa final, la deacilación tioéster permite reconstruir el grupo carboxilo en el sustrato hidrolizado que es concertado con la liberación de una enzima activa (G) [6,8,9].

### 1.3. Proteinasas obtenidas a partir del látex de *Carica papaya L.*

En látex del árbol tropical de *Carica papaya L* está presente una mezcla de endopeptidasas de gran interés científico e industrial, algunas de las propiedades fisicoquímicas de los principales componentes se muestran en la Tabla 1.3. Las cuatro enzimas comparten un plegamiento común característico de CPs de la familia de la papaína, presentan masa molar similar (24 KDa) y puntos isoeléctricos básicos (8.8–11.4), aun cuando su respectiva estructura primaria (secuencia de aminoácidos) presenta considerables diferencias [10].

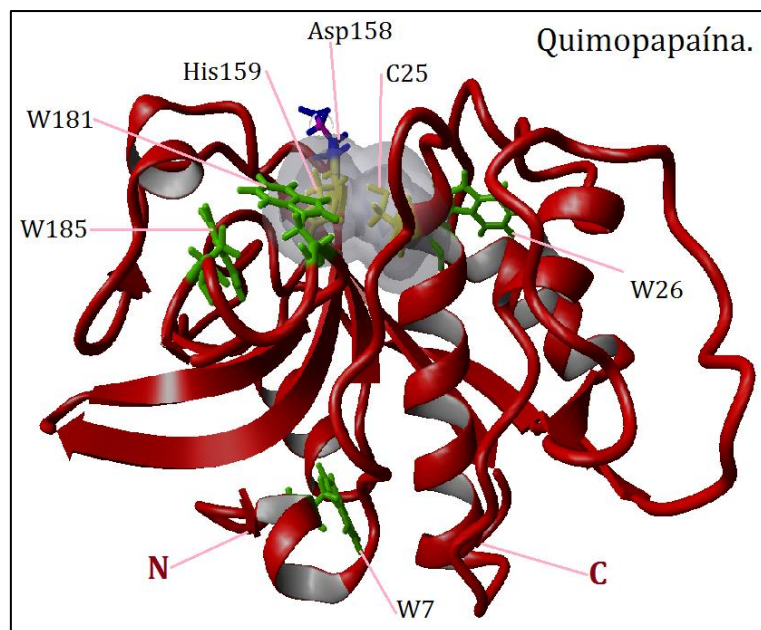
**Tabla 1.3** Propiedades de proteinasas contenidas en látex de *Carica papaya L.*

Proteinasa	Clasificación	pI	Identidad.	Residuos.	M (g/mol)
Papaína	EC 3.4.22.2	8.8		212	23406
Quimopapaína A y B, y otras isoformas	EC 3.4.22.6	10.4	59.40%	218	23636
Glicil-endopeptidasa	EC 3.4.22.25	10.6	67.90%	216	23301
Caricaína	EC 3.4.22.20	11.4	69.80%	216	23264

Donde: pI, punto isoelectrico; Identidad respecto a papaína; M, Masa molar calculada en base a la secuencia de aminoácidos; Datos tomados de Sumner *et al.* (1993)[10].

La quimopapaína es la única entre los miembros de la familia de la papaína que existe como una mezcla constituida principalmente de dos formas, nombradas A y B. Las dos formas son inmunológicamente indistinguibles y presentan idéntica secuencia de aminoácidos [11,12,13]. Sin embargo, difieren en punto isoeléctrico (10.2, forma A; 10.4, forma B), coeficiente de extinción molar ( $\epsilon$ ) a 280 nm (para quimopapaína A,  $\epsilon_{280} = 4.64 \times 10^4 \text{ M}^{-1} \text{ cm}^{-1}$ , quimopapaína B,  $\epsilon_{280} = 5.00 \times 10^4 \text{ M}^{-1} \text{ cm}^{-1}$ ) y en reactividad de sus sitios activos [3,12]. Maes *et al.* (1996) [12] reportaron la estructura cristalina de quimopapaína (Figura 1.3), en base a sus estudios concluye se trata de quimopapaína B.

La quimopapaína tiene propiedades fisicoquímicas muy parecidas a los otros componentes del látex de papaya, lo cual dificulta su purificación. Esta enzima se presenta en múltiples formas (son identificadas cinco formas) que pueden ser separadas mediante cromatografía de intercambio catiónico, las cuales muestran ser inmunológicamente idénticas [3,11,12].



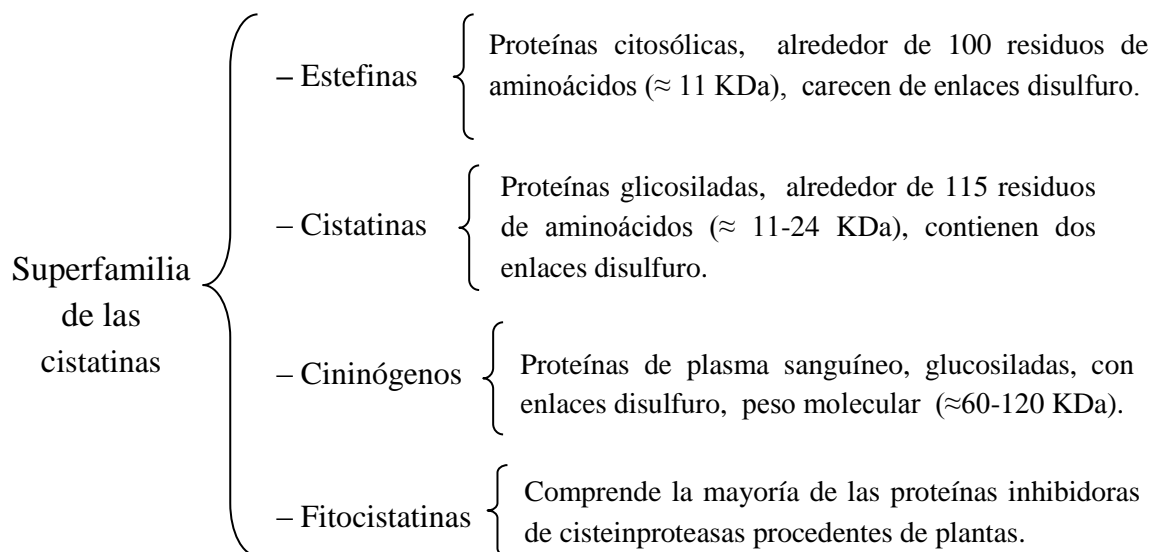
**Figura 1.3** Estructura de quimopapaína (PDB 1YAL dibujada con el programa YASARA [14]). La cadena polipeptídica es plegada en dos dominios de aproximadamente el mismo tamaño pero diferente conformación. Un dominio es constituido principalmente por estructura hélice  $\alpha$  (dominio L: residuos 10-111 y 211-218) y otro esencialmente construido por interacciones de hojas  $\beta$  extendida (dominio R: residuos 1-9 y 112-210). El sitio activo ( $\text{Cys}^{25}$ ) es localizado entre la interface de estos dos dominios [12]. Contiene tres cistinas ( $\text{C}^{22}$ - $\text{C}^{63}$ ,  $\text{C}^{56}$ - $\text{C}^{95}$  y  $\text{C}^{153}$ - $\text{C}^{204}$ ) y dos cisteínas ( $\text{C}^{25}$  y  $\text{C}^{117}$ ) [15].

Buttle & Barret (1984) [11] mostraron que la quimopapaína obtenida a partir de látex fresco de *Carica papaya L* eluye como quimopapaína B y sugieren que las otras formas de quimopapaína son artefactos producidos durante la preparación comercial del látex. Las modificaciones de quimopapaína pueden ser resultado de la actividad de alguna aminopeptidasa o proteinasa permitiendo la hidrólisis de diferentes aminoácidos en el extremo amino terminal (N, ver Figura 1.3) asociados con la preparación de quimopapaína A. Estas modificaciones pueden ser el origen de la diferencia en elución en cromatografía de intercambio catiónico y en los residuos del entorno del sitio activo.

Jacquet *et al.* (1989) [15] apoyaron la teoría de Buttle & Barret (1984) [11] respecto a la presencia de un gen responsable de la expresión de quimopapaína. En sus estudios acerca de la secuenciación de estructura primaria de la quimopapaína, reportaron no observar trazas de microheterogeneidad, lo que sugiriera fuertemente se tratara de una mezcla de varias formas de quimopapaína con secuencia de aminoácidos idéntica [15].

## 1.4. Cistatinas: Inhibidores de cisteinproteasas.

Las cistatinas se encuentran extensamente distribuidas en animales y plantas, de manera general se asume que estos inhibidores están involucrados en la regulación fisiológica y procesos patológicos [16]. Los miembros de la superfamilia de la cistatina han sido agrupados en familias de acuerdo a su localización, tamaño y complejidad de su cadena polipeptídica (Figura 1.4) [4,17].



**Figura 1.4** Clasificación de cistatinas.

En estudios de asociación de CPs con cistatina C humana y cistatina de pollo, hay dos eventos relevantes. El primero, la formación del complejo de cistatina con CPs se lleva a cabo tanto con enzimas en su forma activa, como teniendo su sitio catalítico bloqueado (enzima inactiva), donde se ha observado que la afinidad decrece apreciablemente, en proporción al tamaño del grupo inactivante. Por ejemplo, la carboximetilación de papaína reduce la  $K_d$  con cistatina de pollo alrededor de tres órdenes de magnitud (de 60 fM a 15 pM) [18]. El segundo es la formación del complejo de CPs con cistatinas modificadas en su extremo amino terminal. Por ejemplo, cistatina C cortada en  $G^{11}-G^{12}$  y cistatina de pollo cortada en  $G^9-A^{10}$ , ambas presentan  $K_i$  de tres órdenes de magnitud menor con respecto a cistatina nativa [19].

Las estructuras tridimensionales de varias cistatinas de la familia de cistatinas y estefinas en forma libre o en complejo con una proteasa, han sido determinadas por cristalografía de rayos X o espectroscopía de NMR (Resonancia Magnética Nuclear) y se encuentran

disponibles en el repositorio Protein Data Bank (PDB) [20]. El modelo de interacción con papaína fue propuesto por Bode *et al.* (1988) [21], este modelo fue confirmado por Stubbs *et al.* (1990) [22] con la estructura molecular del complejo papaína–estefina B (PDB 1STF). El mecanismo de inhibición no involucra el residuo catalítico Cys<sup>25</sup>, el cual se encuentra a una distancia muy lejana para ser atacado por el segmento amino terminal del inhibidor [21].

### 1.5. Cistatina de clara de huevo de pollo.

Algunas de las proteínas que constituyen la clara de huevo poseen actividades biológicas, estas incluyen actividad antibacteriana, antihipertensiva, inmunomoduladora, antitumoral y antiviral. También poseen propiedades antiadhesivas, antioxidantes y son inhibidoras de proteasas, poseen biodisponibilidad de nutrientes y lípidos funcionales. Por todo esto, es de gran importancia el estudio de los componentes del huevo en la salud humana, para la prevención y tratamiento de enfermedades [23].

La clara de huevo de pollo es una mezcla compleja que está constituida de proteínas (9.7–10.6%), carbohidratos (0.4–0.9%) y lípidos (0.03%) [23]. Algunas de las propiedades fisicoquímicas de las proteínas que constituyen la clara de huevo se muestran en la Tabla 1.4.

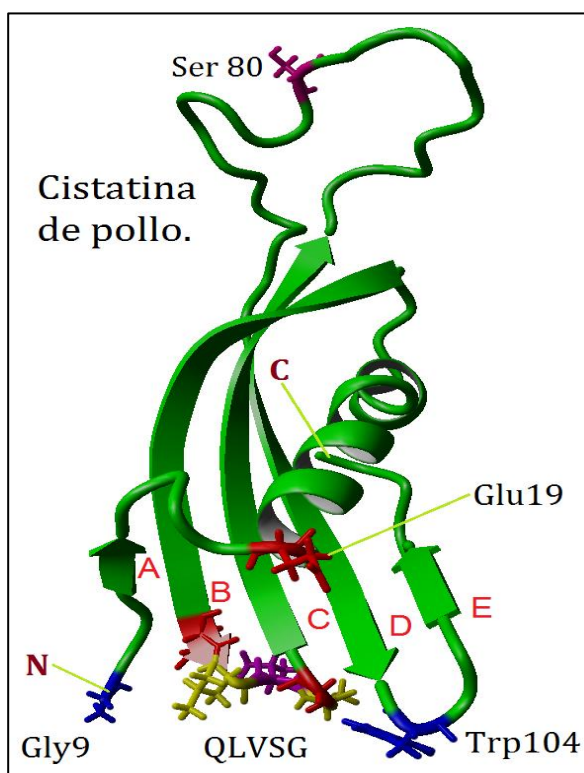
**Tabla 1.4** Propiedades fisicoquímicas de proteínas presentes en clara de huevo.

Proteína	Albumen <sup>1</sup> (% peso seco)	Peso molecular <sup>2</sup> (KDa)	pI <sup>2</sup>	T <sub>d</sub> <sup>2</sup> (°C)
Ovoalbúmina	54	44.5	4.5	84
Ovotransferrina (Conalbúmina)	12	60 – 95	6.0 y 6.6	61
Ovomucoide	11	26.1 – 28.3	3.9 - 4.3	77
Ovomucina	3.5	5.5 – 8.3	4.5 – 5.0	-----
Lisozima	3.4	14.3-14.6	10.7	75
Globulina G <sub>2</sub>	4	30	5.5	92.5
Globulina G <sub>3</sub>	4	45	5.8	-----
Cistatina	0.05	13.2 <sup>3</sup>	6.5 – 5.6	≥120 <sup>4</sup>

Donde: pI, punto Isoeléctrico; T<sub>d</sub>, Temperatura de desnaturalización; <sup>1</sup> Kovacs-Nolan *et al.* (2005) [23]; <sup>2</sup> Carraro *et al.* (2006) [24]; <sup>3</sup> Schwabe *et al.* (1984) [25]; <sup>4</sup> Zerovnik *et al.* (1997) [26].

La cistatina de pollo es un inhibidor de cisteinproteasas que se encuentra en clara de huevo (60  $\mu\text{g/mL}$ ) y suero de pollo (1  $\mu\text{g/mL}$ ) [27], también en pulmón, molleja, cerebro, corazón, oviducto y es expresado por células de músculo y en el desarrollo embrionario de pollo [28,29]. La cistatina de pollo también presenta actividad antimicrobiana, inmunomoduladora y antitumoral e inhibe la degradación de huesos [23].

La cistatina de pollo es el inhibidor mejor caracterizado de la familia de cistatinas, es aislada como una mezcla de dos formas isoelectricas que presentan secuencia de aminoácidos idénticas y no hay evidencia que presenten diferencia en la unión a papaina [30]. Su estructura cristalográfica fue resultada por Bode *et al.* (1988) [21] (ver Figura 1.5), está compuesta de 116 residuos con una masa molecular calculada de 13143 g/mol [25]: La forma 1 (no fosforilada) con punto isoelectrico (pI) de 6.5 y la forma 2 (fosforilada en Ser<sup>80</sup>, P<sub>Ser</sub>) con pI de 5.6. La carga negativa adicional de P<sub>Ser</sub> es responsable de bajar una unidad de pH el pI de la forma 1 [31].



**Figura 1.5** Estructura de cistatina de pollo (PDB 1A67 dibujada con el programa YASARA). La molécula consiste principalmente de estructura beta constituida de cinco hebras antiparalelas (A, B, C, D, E) que envuelven una estructura hélice de cinco giros, además de contener dos puentes disulfuro (C<sup>71</sup>-C<sup>81</sup> y C<sup>95</sup>-C<sup>115</sup>). La forma cortada de cistatina inicia con Gly9 (N) y termina con Q116 (C) [21].

No se tiene conocimiento de si las dos formas de cistatina ocurren de forma natural o si una es formada a partir de la otra durante el almacenamiento del huevo o purificación. Se conoce que P-Ser en péptidos, como en proteínas, es sensible a condiciones alcalinas y es rápidamente destruida por  $\beta$ -eliminación del grupo fosfato. Adicionalmente, la forma 2 puede ser desfosforilada en el huevo, especialmente por fosfatasa alcalina (durante el almacenaje la clara de huevo presenta un pH de 9.5). Las cantidades relativas de la forma 1 y 2 después de la purificación no reflejan las reales *in vivo* [31].

Las dos formas de cistatina de pollo presentan constantes de afinidad por CPs similares. Para enzimas activas, éstas son idénticas considerando el error experimental [18,32] y para enzimas en su forma inactiva (sitio catalítico bloqueado) son indistinguibles una de la otra [30].

## **1.6. Interacción proteína-proteína y su dependencia con el pH.**

Las interacciones electrostáticas son uno de los factores más importantes a ser considerados al analizar la función de moléculas biológicas. La presencia de sitios ionizables, los cuales pueden intercambiar protones con su entorno, provoca el fenómeno de dependencia del pH en proteínas generando una importante influencia en su función de las mismas. Una correcta predicción de los estados de protonación es importante para el análisis de mecanismos de las enzimas, estabilidad de proteínas y reconocimiento molecular [33].

La asociación de proteínas y su dependencia con el pH está fuertemente vinculada con las interacciones electrostáticas (los puentes de hidrógeno, las fuerzas de van der Waals y las interacciones entre carga y carga), la asociación también depende de factores como son las formas estructurales complementarias e interacciones hidrofóbicas [34].

La energía libre de unión puede ser afectada notablemente por la fuerza iónica y el pH de la solución. Por ejemplo, las fuerzas de interacción electrostática de largo alcance son significativamente reducidas a fuerza iónica alta, incluso, en algunos casos estas condiciones permiten apagar las interacciones electrostáticas experimentalmente para estudiar sus efectos sobre la unión proteína-proteína. El pH determina la carga de los residuos ionizables, los cuales a su vez modulan las interacciones electrostáticas que contribuyen a la unión proteína-proteína. Además, estas cargas pueden cambiar debido a la unión (dando origen a la dependencia del pH de la constante de unión) afectando la energía libre de unión [35].



En la actualidad una gran variedad de métodos computacionales han sido desarrollados para estudiar las interacciones electrostáticas en sistemas bimoleculares, los cuales pueden ser clasificados en dos categorías: métodos con solvente explícito, los cuales tratan al solvente en completo detalle atómico y métodos con solvente implícito, que representan al solvente a través de sus efectos promedios sobre el soluto [33].

Una gran variedad de métodos computacionales y algoritmos han sido desarrollados para el estudio de las proteínas y sus interacciones con otras moléculas. Métodos como la simulación con dinámica molecular, simulación con dinámica Browniana, cálculo de  $pK_a$  en proteínas, algoritmos de diseño de proteínas y reconocimiento proteína–proteína son ampliamente utilizados en investigación en Biología Moderna. Todos estos algoritmos emplean estructuras tridimensionales de moléculas de proteína para predecir y analizar características, tales como, actividad catalítica, vía de plegamiento, estabilidad, solubilidad y sitio específico de unión de ligando [36].

La manera más simple de realizar cálculos de la energía de unión sobre un complejo bimolecular es asumir una conformación rígida, es decir, sin algún cambio conformacional sobre la unión, lo cual es claramente no realista, pero frecuentemente proporciona estimaciones iniciales útiles para afinidades relativas de la unión bimolecular [33].

Efectos de la fuerza iónica ( $I$ ) sobre la energía libre de unión ( $\Delta G_b$ ) proteína–proteína han sido incluidos en muchos estudios computacionales, mientras comparativamente pocos estudios han mostrado el correspondiente efecto del pH. La dependencia  $\Delta G_b$  con respecto la variación de pH puede ser muy compleja si están presentes varios grupos en la interfase de la unión proteína–proteína cambiando su estado de protonación, mientras que la dependencia a la fuerza iónica usualmente se describe en función del logaritmo natural de  $I$ , logrando exitosamente predecir la magnitud de la correlación, la cual depende de la forma y distribución de cargas de las moléculas [35].

En el campo de la medicina, los estudios de sistemas de interacción enzima–inhibidor (unión proteína–proteína) contribuyen a comprender el mecanismo de la patofisiología de enfermedades, así como, ayudar en la identificación y priorización de blancos terapéuticos. La realización de estos estudios permiten generar nuevas hipótesis para el diseño de mejores fármacos para blancos específicos, o asociados a enfermedades (por ejemplo, diferentes tipos de cáncer), incluso pueden dar la pauta a descubrir nuevos fármacos.

El gran interés de estudiar la interacción entre quimopapaína–cistatina de pollo como modelo enzima–inhibidor radica en su gran potencial farmacéutico y biotecnológico, ambas proteínas son termoestables y presentan una gran estabilidad estructural y funcional a valores de pH extremos.

En el presente trabajo con el propósito de elucidar y comprender el mecanismo de interacción del sistema en estudio, se estudió la dependencia de la constante de unión ( $K_b$ ) con respecto la variación de pH por métodos espectroscópicos, y se analizó la interacción con métodos computacionales empleando un modelo molecular de la estructura del complejo quimopapaína–cistatina de pollo.

### **1.7. Importancia y aplicaciones del estudio de cisteinproteasas y sus inhibidores.**

La investigación y estudio acerca de proteasas y sus inhibidores es efectuada con el firme propósito de conocer más aplicaciones potenciales en medicina, agricultura y biotecnología. Para ello se han efectuado y se continúan realizando estudios *in vivo* e *in vitro* de la formación de complejos proteasa–inhibidor.

Diversos estudios *in vivo* han demostrado que proteínas inhibidoras de cisteinproteasas endógenas regulan endopeptidasas cisteínicas lisosomales (por ejemplo: Catepsina B, H, L, S y X). Se encuentran reportados estudios de estos inhibidores en el tratamiento de cáncer y metástasis, pancreatitis, artritis, arteriosclerosis, diabetes, gingivitis periodontal, infecciones, inflamación, trastornos en el hígado, pulmón, corazón, riñón y músculos, así como en la enfermedad de Alzheimer y de Batten [37].

Con la finalidad de explicar el mecanismo de inhibición de cisteinproteasas por cistatinas se han realizado numerosos estudios espectroscópicos, cinéticos y cristalográficos. El tipo de inhibición que presenta este sistema se denomina mecanismo de exosítio, en el cual la unión se realiza en la zona de la superficie adyacente a su centro activo de la proteinasa impidiendo el acceso de la molécula de sustrato al sitio activo, sin bloquear directamente el residuo catalítico (Cys) de la enzima [38]. Se ha mostrado que estas enzimas con su sitio activo bloqueado pueden unirse a cistatinas, presentando una afinidad más baja [6,18,39]. A pesar de la homología estructural y modos de inhibición similar, la cistatina presenta diferentes afinidades con las diversas enzimas de la familia de la papaína del orden menor a 0.1 nM [6] (ver Tabla 1.5).

**Tabla 1.5** Constantes de inhibición ( $K_i$ ) y disociación ( $K_d$ ) de sistemas cisteinproteasa–cistatina.

Enzima		Estefina A	Estefina B	Cistatina C	Cistatina de pollo.
Papaína	$K_i$	0.019 <sup>a</sup>	0.12 <sup>a</sup>	0.005 <sup>a</sup>	0.005 <sup>a</sup>
	$K_d$			0.000011 <sup>c</sup>	0.00006 <sup>d</sup>
Catepsina B	$K_i$	8.2 <sup>a</sup>	73 <sup>a</sup>	0.25 <sup>a</sup>	1.7 <sup>a</sup>
Catepsina H	$K_i$	0.31 <sup>a</sup>	0.58 <sup>a</sup>	0.28 <sup>a</sup>	0.064 <sup>a</sup>
Catepsina L	$K_i$	1.3 <sup>a</sup>	0.23 <sup>a</sup>	0.005 <sup>a</sup>	0.019 <sup>a</sup>
Catepsina X	$K_i$	1.7 <sup>b</sup>	>250 <sup>b</sup>	12 <sup>b</sup>	15 <sup>b</sup>
Quimopapaína	$K_d$				0.0009 <sup>e</sup>

**Donde:**  $K_i$ , constante de inhibición (nM);  $K_d$ , constante de disociación (nM);

<sup>a</sup> Barrett (1987) [40]; <sup>b</sup> Klemenčič *et al.* (2000) [41]; <sup>c</sup> Lindah *et al.* (1992) [42];

<sup>d</sup> Björk *et al.* (1989) [43]; <sup>e</sup> Björk & Ylinenjärvi (1990) [39]; <sup>f</sup> calculada a partir de las constantes de velocidad de asociación y disociación.

Como modelo de estudio para este tipo de sistemas enzima–inhibidor, en el presente trabajo, se estudió la dependencia de la constante de la unión ( $K_b$ ) con respecto a la variación del pH de la cisteinproteasa S-carboximetil-quimopapaína (SCM-quimopapaína) con cistatina de clara de huevo de pollo, que es un inhibidor de cisteinproteasas.

## 2. OBJETIVOS.

### 2.1. Objetivos Generales.

- A) Estudiar la dependencia de la constante de unión ( $K_b$ ) con la variación de pH para la asociación de la cisteinproteasa S-carboximetil-quimopapaína (SCM-quimopapaína) con el inhibidor cistatina de pollo (CEW), en presencia de amortiguadores de pH a dos fuerzas iónicas de NaCl (0.1 M y 0.1M) 25 °C.
- B) Caracterizar por métodos espectroscópicos y computacionales la asociación del sistema quimopapaína–cistatina de pollo empleando un modelo molecular de la estructura del complejo enzima–inhibidor.

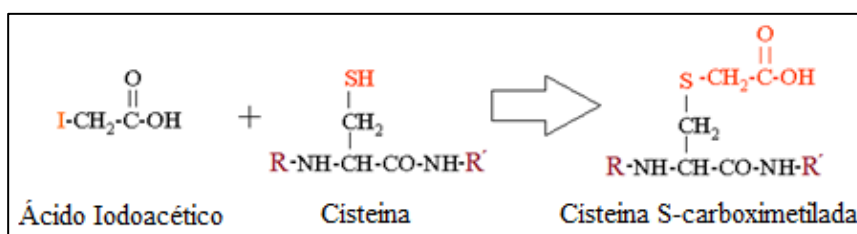
### 2.2. Objetivos Particulares.

- a) Purificar SCM-quimopapaína a partir de un extracto comercial parcialmente purificada marca SIGMA.
- b) Purificar cistatina de pollo a partir de clara de huevo de pollo, establecer el método de purificación y caracterizar la proteína (propiedades espectrales, actividad inhibitoria de cisteinproteasas y pruebas de pureza).
- c) Estudiar la dependencia de la variación del pH en el intervalo de [3.5, 10.5] sobre la constante de unión del sistema SCM-quimopapaína–cistatina de pollo, determinando la constante de disociación ( $K_d$ ) por medio del método de titulación fluorimétrica a una de temperatura de 25 °C en presencia de NaCl 0.1 M y 0.01 M.
- d) Examinar la interacción de SCM-quimopapaína y cistatina de pollo con la construcción de un modelo molecular. Analizar los grupos ionizables que se encuentran en la zona de interacción de las moléculas y su efecto en la dependencia de la constante de unión con el pH.

### 3. Materiales y Métodos.

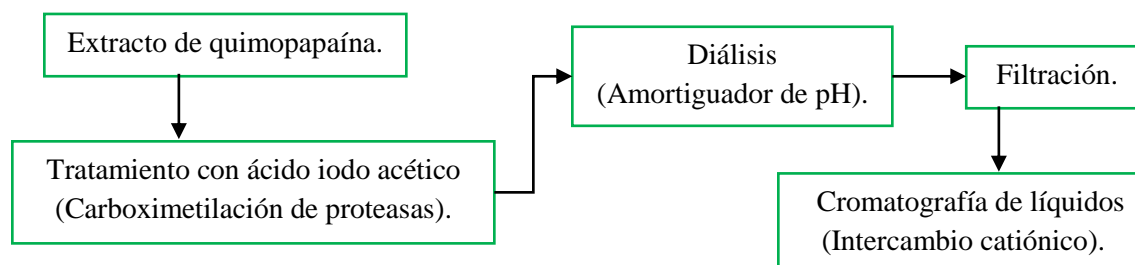
#### 3.1. Método de purificación de SCM-quimopapaína.

La quimopapaína fue purificada a partir de un extracto de quimopapaína comercial. Para evitar autohidrólisis, su sitio activo (Cys<sup>25</sup>) fue bloqueado previamente, esto se logró carboximetilando las cisteínas libres Cys<sup>25</sup> y Cys<sup>117</sup> de quimopapaína (Figura 3.1).



**Figura 3.1** Carboximetilación (SCM-) de cisteína con ácido iodo acético.

La purificación de quimopapaína se efectuó mediante cromatografía de intercambio catiónico aplicando un gradiente bifásico de incremento en la fuerza iónica del amortiguador de pH [44]. Este procedimiento de purificación (Figura 3.2) es una modificación al método previamente reportado por Solis-Mendiola *et al.* (1992) [44].



**Figura 3.2** Procedimiento de purificación de SCM-quimopapaína.

La carboximetilación de las proteasas contenidas en el extracto de quimopapaína se efectuó de la siguiente manera. Se pesaron 200 mg de extracto de quimopapaína (marca SIGMA, lote:124F80751) y se disolvieron en 80 mL de buffer fosfatos 0.1 M, pH 7.0. Las proteasas sulfhidricas fueron activadas con cisteína 2 mM y se mantuvo en agitación por 10 min a temperatura ambiente. Una vez efectuada esta operación fue adicionado ácido iodoacético hasta una concentración final de 10 mM. La solución se mantuvo en agitación constante a temperatura ambiente por una hora. La solución resultante fue dializada contra amortiguador de fosfatos 50 mM, pH 7.0 y filtrada a través de una membrana de 0.45  $\mu\text{m}$  (Durapore, HV).

La solución resultante se cargó a una columna de intercambio catiónico Mono S 5/50 GL (GE Healthcare) montada en un equipo FPLC (Fast Performance liquid chromatography, GE ÄKTApriime plus), equilibrada con amortiguador de fosfato de sodio 50 mM, pH 7.0. La desorción se realizó con un gradiente bifásico de NaCl, de 0 a 0.4 M en un volumen de elución ( $V_e$ ) de 15 mL y de 0.4 a 0.8 M en  $V_e$  de 10 mL, a un flujo de 0.7 mL/min.

La fracción correspondiente a S-carboximetil-quimopapaína A (SCM-quimopapaína) identificada en base a comparación con cromatogramas de trabajos previos [11,44], fue colectada y dializada contra amortiguador de fosfatos 10 mM, pH 7.4.

Posteriormente, esta solución fue concentrada mediante ultrafiltración, empleando una membrana de corte molecular de 10 KDa (Amicon Bioseparations), a una concentración cercana a 1 mg/mL. La concentración de SCM-quimopapaína fue calculada con base en su  $A_{280, 1\text{cm}}^{1\%} = 18.2$  [45], realizando una corrección por dispersión de luz [46].

### **3.1.1. Pruebas de pureza de SCM-quimopapaína.**

Se realizaron dos pruebas de pureza a la SCM-quimopapaína, se verificó su homogeneidad mediante recromatografía de intercambio catiónico y se efectuó el análisis con espectrometría de masas.

#### *A. Recromatografía de intercambio catiónico.*

La recromatografía de SCM-quimopapaína se realizó mediante cromatografía de intercambio catiónico (columna Mono S 5/50 GL) empleando amortiguador de fosfato 50 mM, pH 7.0 y un gradiente bifásico como se indican en Sección 3.1.

#### *B. Espectrometría de masas.*

Una muestra de SCM-quimopapaína purificada se envió al Laboratorio Divisional de Espectrometría de Masas (LDEM), ubicado en la UAM-Iztapalapa.

La muestra fue analizada por espectrometría de masas MALDI-TOF, en un equipo MALDI-TOF/TOF AutoFlex™ Bruker. MALDI por sus siglas en inglés Matrix Assisted Laser Desorption/Ionization (Desorción/Ionización Láser Asistida por Matriz) y TOF por el detector de iones acoplado al MALDI y cuyo nombre procede también de sus siglas en inglés Time-Of-Flight (Tiempo de Vuelo).

La espectrometría de masas es una técnica que permite determinar las masas atómicas o moleculares. MALDI-TOF es una técnica de ionización suave que permite el análisis de biomoléculas (por ejemplo biopolímeros como las proteínas, péptidos, azúcares y lípidos) y moléculas orgánicas grandes (por ejemplo polímeros, dendrímeros) que podrían ser susceptibles de fragmentación al ser sometidas a otros métodos de ionización más convencionales.

La preparación de la muestra de cistatina se realizó eliminando los iones presentes del amortiguador de pH, dializando la muestra contra agua desionizada, finalmente se concentró a 2.4 mg/mL de proteína y se entregó al LDEM.

La muestra de SCM-quimopapaína dializada fue mezclada en proporción adecuada (1:50,000, aproximadamente) con la matriz HCCA ( $\alpha$ -cyano-4-hydroxycinnamic acid) y las muestras preparadas mediante la técnica Dried Droplet aplicadas sobre una placa Ground Steel. Se obtuvieron los espectros de masa en modo lineal positivo utilizando los métodos LP\_5-20\_kDa. Previamente, el espectrómetro de masas fue calibrado externamente con estándares de citocromo C, (M+H)<sup>+</sup> 12,361.96 Da, apomioglobina, (M+H)<sup>+</sup> 16,952.27 Da, y albúmina, (M+H)<sup>+</sup> 66,429.09 Da.

### **3.1.2. Propiedades espectrales de quimopapaína purificada: Espectros de fluorescencia y dicroísmo circular (DC).**

#### *A. Espectros de fluorescencia.*

Los espectros de emisión de fluorescencia se obtuvieron en un espectrofluorómetro ISS K2 (Urbana USA) equipado con un sistema de control de temperatura tipo peltier. Se utilizó una celda de 1 cm de paso óptico, al efectuar las mediciones las muestras se mantuvieron en agitación magnética constante, a una temperatura de 25 °C. En todas las mediciones se utilizó una rejilla de emisión con un ancho de banda de 8 nm y una de excitación de 16 nm.

Los espectros de emisión de fluorescencia se registraron de 300 nm a 400 nm, con una longitud de onda ( $\lambda$ ) de excitación de 280 nm. En todos los ensayos, se reportan los espectros de emisión de fluorescencia normalizados con respecto al valor máximo registrado para quimopapaína, previamente restando la señal del amortiguador de pH. La concentración de SCM-quimopapaína fue de 0.1  $\mu$ M en presencia de amortiguador de Tris-base 50 mM, NaCl 0.1 M, pH 7.4.

## B. Espectros de DC.

Los espectros de dicroísmo circular (DC) en la región del UV-lejano de 190 a 250 nm se obtuvieron en un espectropolarímetro JASCO J-715 (Jasco Inc., Easton, MD) equipado con un controlador de temperatura tipo peltier PTC-348WI. Los experimentos se realizaron a una temperatura de 25 °C, empleando amortiguador de fosfato de sodio 50 mM, pH 7.0. La medición se efectuó a una concentración de 0.1 mg/mL en una celda de 0.1 cm de paso óptico. En el software del equipo se indicó obtener el promedio de tres espectros. Para el tratamiento de los datos se restó la señal del amortiguador de pH y el espectro correspondiente se presenta en elipticidad molar por residuo medio,  $[\theta]_{MRW}$  (Ecuación 3.1):

---

Ecuación 3.1 Elipticidad molar por residuo medio,  $[\theta]_{MRW}$  [47].

---

$$[\theta]_{MRW} = \frac{\theta_{obs} \times 100 \times MRW}{C \times l}$$

Donde:  $\theta_{obs}$ , la señal de DC en grados;  $MRW$ , masa de residuo medio (para quimopapaína es 108.4 g/mol);  $l$ , paso óptico de la celda en cm;  $C$ , es la concentración en g/mL.

---

## 3.2. Método de purificación de cistatina de pollo (CEW).

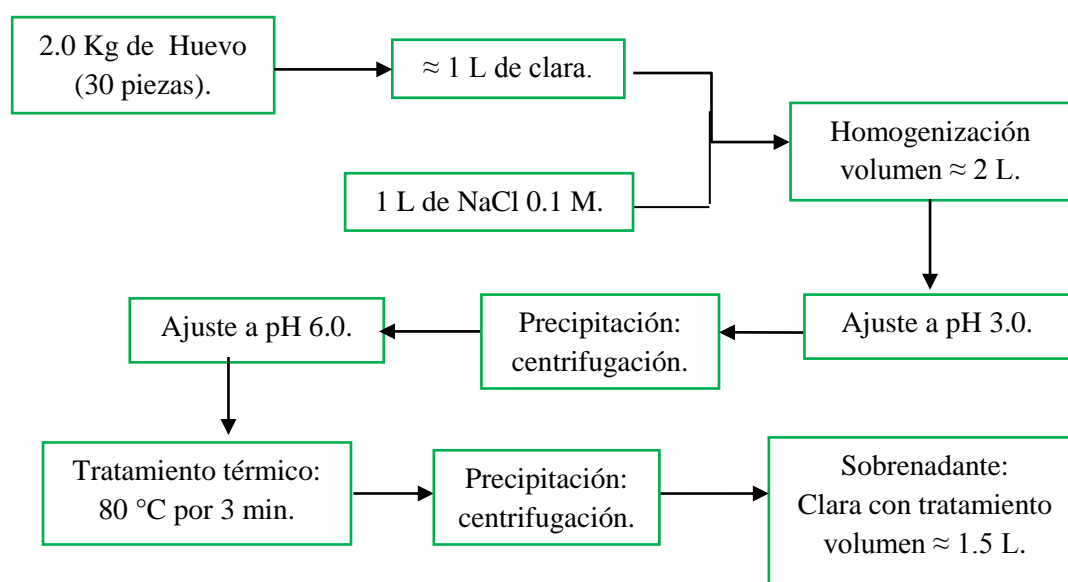
La cistatina de pollo (CEW) fue purificada a partir de clara de huevo mediante la modificación de métodos previamente reportados por otros grupos de investigación [16,29,30,48-51] (ver Apéndice A).

La CEW está presente en un 0.05% del contenido de proteína en clara de huevo [23], con una concentración de 60  $\mu$ g/mL [27], se ha reportado ser estable a pH extremos [8] y temperaturas altas [52]. Aprovechando estas propiedades, el método de purificación de CEW desarrollado en el presente trabajo involucra un tratamiento de pH y temperatura a la clara de huevo, continuando con una cromatografía de afinidad y seguida de cromatografía de exclusión molecular.



### 3.2.1. Preparación de clara de huevo.

El tratamiento que se le dio a la clara de huevo es descrito en el esquema de la Figura 3.3, el procedimiento fue el siguiente: La clara de huevo fue homogenizada y diluida con un volumen igual de solución NaCl 0.1 M. Después de esta operación, se ajustó el pH a 3.0 con HCl 3.0 M, manteniendo la solución en agitación constante por una hora a 4 °C. La proteína insoluble fue removida por precipitación, centrifugando a 4000g por 30 min. Al sobrenadante colectado se le ajustó el pH a 6.0 con NaOH 3.0 M y permaneció en agitación constante por una hora. La solución resultante se calentó a 80 °C por tres minutos y se enfrió a temperatura ambiente. La proteína insoluble fue removida por precipitación (centrifugando a 4000g por 10 min) y ajustado su pH a 7.0. La concentración de CEW contenida en el sobrenadante de clara con tratamiento es de  $\approx 25 \mu\text{g/mL}$ , la cual es sometida a cromatografía de afinidad.



**Figura 3.3** Tratamiento de clara de huevo.

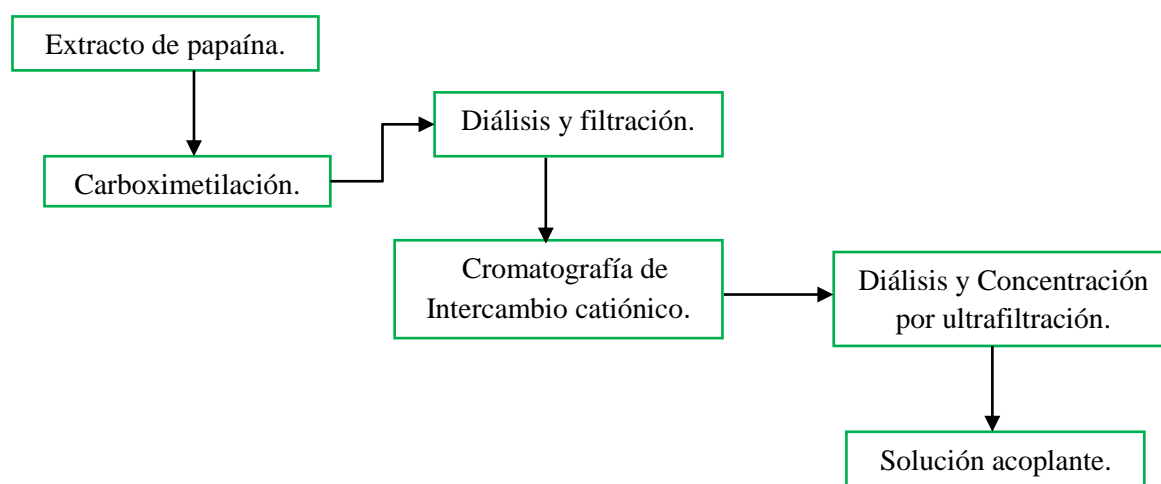
### 3.2.2. Inmovilización de SCM-papaína sobre sefarosa 4B para cromatografía de afinidad.

La inmovilización de enzimas es un proceso en el que se confina o localiza a la enzima en una región definida, por ejemplo un soporte. Empleando el método de inmovilización covalente con bromuro cianógeno (CNBr) activado en sefarosa (4B-CNBr) se inmovilizó S-carboximetil-papaína (SCM-papaína).

La solución de acoplante se preparó de la siguiente manera (Figura 3.4). Se pesaron 200 mg de extracto de papaína dos veces cristalizada (SIGMA, Lote: 43H7025) y se efectuó el proceso de carboximetilación del sitio activo como se describe en la Sección 3.1, empleando amortiguador de fosfatos 0.1 M, pH 6.0.

La purificación de SCM-papaína se efectuó mediante cromatografía de intercambio catiónico como se describe en la Sección 3.1, utilizando amortiguador de fosfato de sodio 0.05 M, pH 6.0 y fosfato de sodio 0.05 M, NaCl 1.0 M pH 6.0 como soluciones eluyentes.

Las fracciones de SCM-papaína se concentraron y dializaron mediante ultrafiltración en presencia de amortiguador carbonato de sodio 0.1 M, NaCl 0.5 M, pH 8.3. La solución resultante contiene 18 mg de SCM-papaína y es la solución acoplante para continuar con el proceso de inmovilización. La cuantificación de SCM-papaína purificada se realizó con base en su  $A^{1\%}_{280\text{nm}, 1\text{cm}} = 25$  (Especificación de proveedor SIGMA).



**Figura 3.4** Preparación de SCM-papaína (solución acoplante).

La preparación de sefarosa se realizó como lo indica el protocolo del proveedor. Se hidrataron 10 gramos de sefarosa 4B-CNBr liofilizada (SIGMA-ALDRICH, lote: 124k1654) con 25 mL de HCl 1 mM (volumen final  $\cong$  35 mL). El gel fue lavado por filtración al vacío con 1000 mL de HCl 1 mM, en un tiempo menor a 15 min.

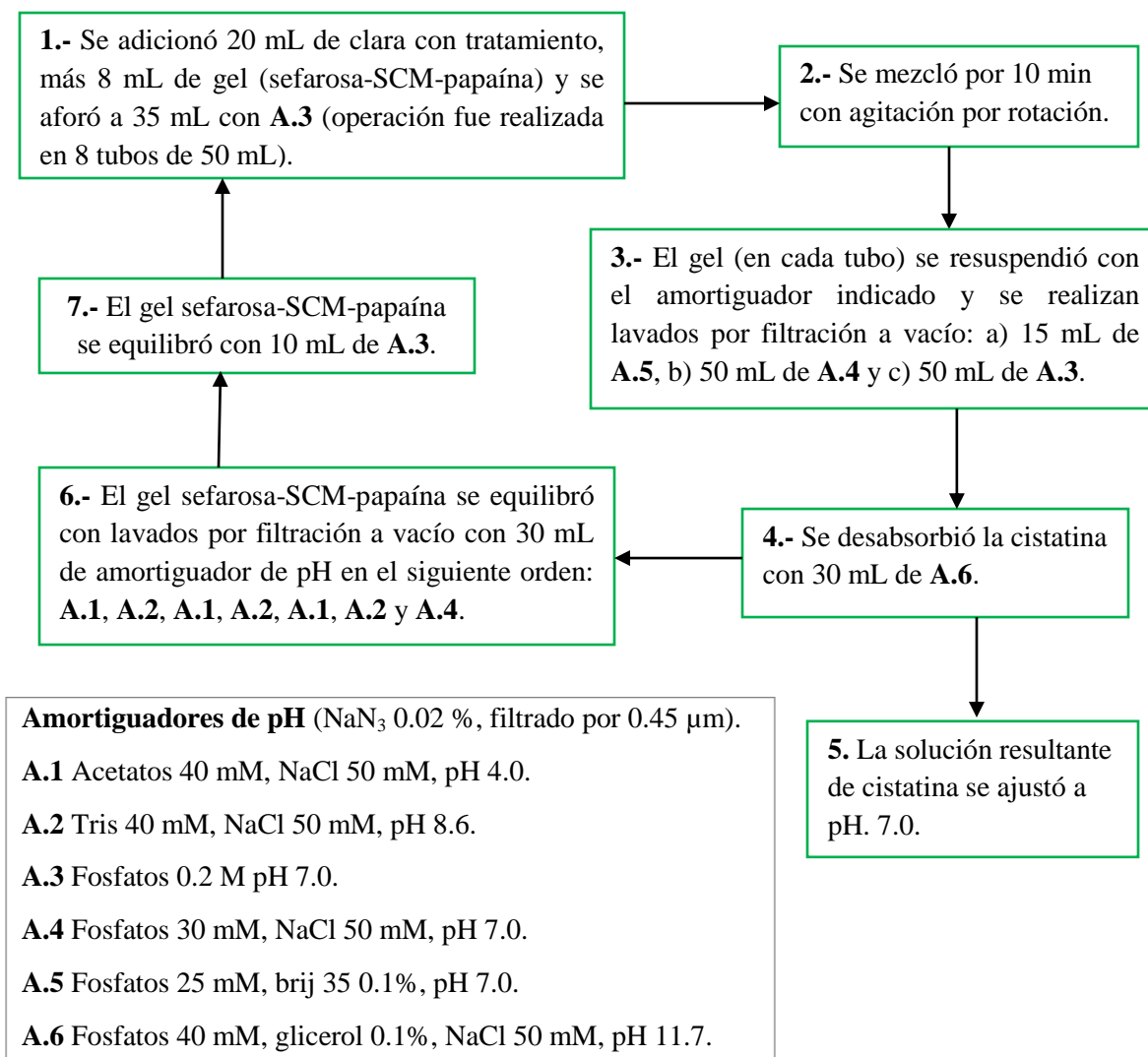
La solución acoplante y la sefarosa fueron mezcladas agitando constantemente mediante rotación por una hora a temperatura ambiente. Una vez pasado este tiempo la mezcla fue filtrada. Para verificar que la proteína fue ligada se registró la absorbancia del filtrado a 280 nm y se comparó con la lectura de la solución acoplante.

La SCM-papaína no ligada de forma covalente (adherido en la superficie del gel) fue eliminada efectuando un lavado con 200 mL de amortiguador  $\text{NaHCO}_3$  0.1 M, NaCl 0.5 M, pH 8.3. Para bloquear los grupos activos remanentes, el gel se resuspendió en 100 mL de amortiguador Tris-HCl 0.1 M, pH 8.0, por 2 horas con ligera agitación por rotación, pasado este tiempo el gel se filtró.

El gel de sefarosa-SCM-papaína fue sometido a tres ciclos de lavado por filtración, alternando el pH de 8.0 a 4.0. Para cada ciclo, se empleó un lavado con 100 mL de amortiguador acetato de sodio 0.1 M, NaCl 0.5 M, pH 4.0, seguida por un lavado de amortiguador Tris-HCl 0.1 M, NaCl 0.5 M, pH 8.0. Finalmente se equilibró el gel de sefarosa-SCM-papaína, mediante un lavado por filtración con 100 mL de amortiguador fosfato de sodio 50 mM, NaCl 0.5 M, Brij 35 1.0 %, pH 6.5.

### **3.2.3. Cromatografía de afinidad para la purificación de CEW.**

La cromatografía de afinidad se efectuó por cromatografía en *batch*, ver Figura 3.5. La determinación de proteína ligada al gel de sefarosa reportada es de 0.6 mg de papaína/g de gel [29]. La SCM-papaína inmovilizada (en 10 g de sefarosa) tiene una capacidad de retención teórica de cistatina de  $\cong$  3 mg. El gel de sefarosa-SCM-papaína se equilibró con amortiguador de fosfato de sodio 50 mM, pH 7.0 y se mezcló con la solución de clara con tratamiento. Se hicieron tres lavados por filtración a vacío con amortiguador de fosfatos de sodio 50 mM, Brij 35 0.1 %, pH 6.5. La CEW ligada a sefarosa-SCM-papaína fue desorbida con amortiguador de fosfatos de sodio 50 mM, NaCl 50 mM, glicerol 1 %, pH 11.5. La solución de cistatina obtenida fue ajustada a pH 7.0 y concentrada.



**Figura 3.5** Cromatografía de afinidad para la purificación de CEW (Cromatografía en *batch*).

### 3.2.4. Cromatografía de exclusión molecular para la purificación de CEW.

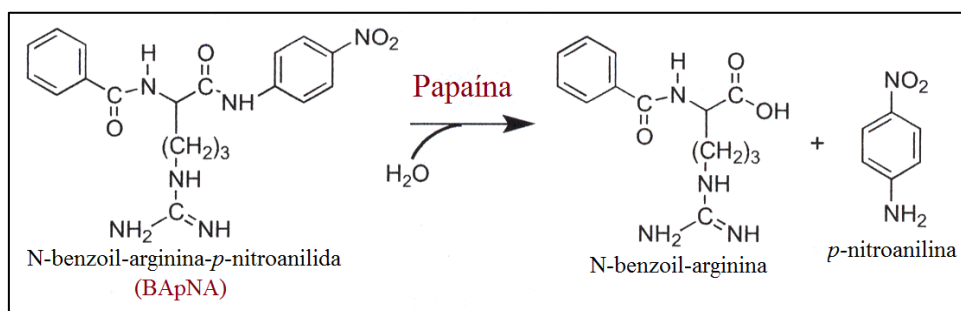
La solución de cistatina resultante de la cromatografía de afinidad fue dializada con amortiguador de fosfato de sodio 30 mM, NaCl 50mM, pH 7.0 y cargada en una columna grado preparativo Superdex 200 HiLoad 16/60 GP (capacidad de separación de 10 a 600 KDa). La fracción de interés se colectó y cargó en una columna Superdex 200 10/300 GL. La fracción identificada como CEW se concentró y dializó con amortiguador de fosfato de sodio 10 mM, pH 7.0. La concentración de CEW fue calculada con base en su  $A_{280\text{ nm}, 1\text{ cm}}^{1\%} = 8.7$  [53], realizando la corrección por dispersión de luz [46].

### 3.2.5. Ensayos de actividad inhibitoria y pruebas de pureza de CEW.

El ensayo de actividad inhibitoria consiste en probar de manera cualitativa que en presencia de CEW hay inhibición de la actividad hidrolítica de papaína. También se realizaron dos pruebas de pureza: electroforesis en gel de poliacrilamida en presencia de dodecil- sulfato de sodio y se efectuó el análisis con espectrometría de masas.

#### A. Ensayo de actividad inhibitoria de CEW.

Se probó la actividad inhibitoria de CEW modificando el método colorimétrico descrito por Barret (1981) [27], el cual se basa en la inhibición de la actividad peptidasa de papaína sobre el sustrato N-benzoil-arginina-p-nitroanilida (BAPNA, marca SIGMA lote:119C-0415), Figura 3.6.



**Figura 3.6** Representación de la hidrólisis del sustrato BAPNA por papaína.

La papaína fue incubada en un volumen de reacción de 870  $\mu$ L, a una concentración de 4.6 a 12  $\mu$ M y activada con cisteína 0.5 mM, en presencia de amortiguador Tris-base 20 mM, pH 8.0, a una temperatura de 40  $^{\circ}$ C. Se preparó una solución stock de sustrato BAPNA a una concentración de 3 mM disuelto en dimetil-sulfoxido (DMSO), la concentración final de sustrato en la celda de reacción fue de 0.1 mM y de DMSO del 3% (v/v).

La cinética de reacción se registró a una longitud de onda de 410 nm, por 10 minutos, empleando un espectrofotómetro ultravioleta-visible (Hewlett Packard, modelo 8453) equipado con control de temperatura sistema peltier (Hewlett Packard, modelo 89090, empleando una celda semimicro de 1 cm de paso óptico.

*B. Electroforesis en gel de poliacrilamida en presencia de dodecil - sulfato de sodio (SDS-PAGE).*

Se realizó SDS-PAGE en condiciones desnaturalizantes empleando un sistema de amortiguador de pH discontinuo (Laemmli), bajo las especificaciones del proveedor BioRad (Número de Catálogo 165-8000). En este sistema las proteínas son desnaturalizadas por calor en amortiguador conteniendo dodecil-sulfato de sodio (SDS) y 2-mercaptoetanol como agente reductor. Se empleó un gel de acrilamida bifásico, para el gel concentrador 4% (amortiguador Trizma 0.5 M, pH 6.8) y para el gel separador 12% (amortiguador Trizma 1.5 M, pH 8.8), como amortiguador de corrida (en la cámara del electrodo) se utilizó Trizma 25 mM, glicina 19 mM, SDS 1%. La electroforesis se realizó aplicando un voltaje de 180 volts.

*C. Espectrometría de masas.*

La preparación de la muestra de cistatina purificada para analizarla mediante espectrometría de masas se realizó como se indica en Sección 3.1.1.B. La muestra se concentró a 1.8 mg/mL de proteína y se entregó al LDEM. Los espectros de masa se obtuvieron en modo lineal positivo utilizando el método LP\_20\_40\_kDa y empleando los estándares señalados en Sección 3.1.1.B.

### **3.2.6. Propiedades espectrales de CEW purificada: Espectros de fluorescencia y dicroísmo circular (DC).**

*A. Espectros de fluorescencia.*

Los detalles del equipo y condiciones de operación para la obtención de los espectros de fluorescencia se indican en la Sección 3.1.2.A. La concentración de CEW fue de 0.12  $\mu$ M en presencia de amortiguador Tris-base 50 mM, NaCl 0.1 M, pH 7.4.

*B. Espectros de DC.*

Se obtuvieron los espectros de dicroísmo circular como se indica en la Sección 3.1.2.B, usando una masa de residuo medio (MRW) de 113.3 g/mol.

### 3.3. Métodos espectroscópicos.

#### 3.3.1. Espectros de absorbancia en la región UV-Visible y dicroísmo circular (DC) en la región UV-lejano en el intervalo de pH de 3.5 a 10.0, a 25 °C.

Los espectros de absorbancia para SCM-quimopapaína y CEW se obtuvieron en un espectrofotómetro ultravioleta-visible (ver detalles del equipo en Sección 3.2.5.A) empleando una celda de 1 cm de paso óptico. Las mediciones se realizaron en la región de 250 a 350 nm, a una concentración de 0.3 mg/mL y 25°C.

Se obtuvieron los espectros de dicroísmo circular (DC) para SCM-quimopapaína y CEW en la región de UV-lejano de 190 a 250 nm, los detalles instrumentales y tratamiento de datos se efectuaron como se indican en la Sección 3.1.2.B. Las mediciones se realizaron a una concentración de 0.1 mg/mL a una temperatura de 25°C, en el tratamiento de datos se empleó una MRW de 108.4 g/mol para quimopapaína y 113.4 g/mol para CEW.

Las mediciones de absorbancia y DC se realizaron en presencia de los siguientes amortiguadores de pH a una concentración de 50 mM: Formiato de sodio (pH 3.5), acetato de sodio (pH 4.5), MES (pH 5.5), fosfato de sodio (pH 6.5, 7.5), Tris-base (pH 8.0, 8.5) y glicina (pH 9.0, 9.5, 10.0).

#### 3.3.2. Determinación de la constante de disociación ( $K_d$ ) por medio del método de titulación fluorimétrica del sistema SCM-quimopapaína–CEW.

La  $K_d$  del sistema SCM-quimopapaína–CEW se determinó mediante el método de titulación fluorimétrica. Los valores registrados de intensidad de emisión de fluorescencia que acompañan la interacción de SCM-quimopapaína y CEW en la titulación fueron evaluados en el modelo matemático de unión bimolecular (ecuación 3.2). La variable independiente es  $\frac{[C]}{[P]}$  y las variables obtenidas experimentalmente son  $F_P$ ,  $F_C$  y  $F_M$ . Realizando una regresión no lineal, se determinó el valor de los parámetros  $K_d$  y  $a$  (constante de apagamiento), usando el programa Origin Pro 8.6.

---

**Ecuación 3.2** Modelo matemático de unión bimolecular para determinar  $K_d$ .

La deducción de esta expresión es presentada en Apéndice C.

---

$$Y = \frac{F_P + F_C - F_M}{F_P} = \frac{a}{2} \left\{ \left( 1 + \frac{K_d}{[P]} + \frac{[C]}{[P]} \right) - \sqrt{\left[ \left( 1 + \frac{K_d}{[P]} + \frac{[C]}{[P]} \right)^2 - 4 \frac{[C]}{[P]} \right]} \right\}$$

Donde:  $F_P$  , fluorescencia de SCM-quimopapaína libre (considerando efectos de dilución);  $F_C$  , fluorescencia de CEW libre (titulación de amortiguador de pH);  $F_M$ , fluorescencia de la mezcla de SCM-quimopapaína y CEW;  $a$  , parámetro que indica el valor asintótico  $Y$  cuando se alcanza la saturación de la enzima;  $K_d$ , constante de disociación en el equilibrio;  $[P]$ , concentración de SCM-quimopapaína total en la celda reacción;  $[C]$ , concentración de CEW total en la celda reacción.

---

Los experimentos de fluorescencia se realizaron en un espectrofluorómetro ISS K2 (ver detalles en Sección 3.1.2.A).

Los espectros de emisión de fluorescencia se registraron de 300 nm a 400 nm, con una longitud de onda ( $\lambda$ ) de excitación de 280 nm. Las concentraciones de SCQ-quimopapaína y CEW fueron de 0.1  $\mu$ M y 0.12  $\mu$ M, respectivamente.

La titulación fluorimétrica se realizó adicionando volúmenes sucesivos (2–5  $\mu$ L) de CEW con una concentración de 9–11  $\mu$ M a 2 mL de una solución de SCM-quimopapaína 0.1  $\mu$ M, de manera paralela esta operación se realizó titulando 2 mL de amortiguador de pH hasta la misma concentración final de CEW. Las mediciones de intensidad de emisión fluorescencia se efectuaron a la longitud donde se presenta el máximo de emisión de la diferencia de los espectros de las proteínas libres y de la mezcla (  $F_P + F_C - F_M$  ), para cada valor de pH, aplicando una  $\lambda$  de excitación de 280 nm.

Las titulaciones se realizaron en los siguientes amortiguadores de pH 50 mM, NaCl 0.1 M: MES (pH 6.0), PIPES (pH 6.5), fosfato de sodio (pH 7.0, 7.4), Tris base (pH 7.2, 7.4, 7.6, 7.8, 8.0, 8.5, 8.8), borato de sodio pH (9.0) y glicina (pH 9.0, 9.2, 9.5, 10.0). También se realizaron titulaciones con fuerza iónica de 0.01 M en los siguientes amortiguadores: formiato de sodio 26 mM (pH 3.5), acetato de sodio 26 mM (pH 4.5), MES 60 mM (pH 5.5), MES 26 mM (pH 6.0) y MES 16 mM (pH 6.5).



### 3.4. Métodos computacionales.

#### 3.4.1. Modelado de la interacción quimopapaína con CEW y simulaciones con Dinámica Molecular (MD).

Se construyó un modelo atómico mediante acoplamiento molecular (*docking*) partiendo de la estructuras de las proteínas libres (quimopapaína: PDB código 1YAL; cistatina de pollo: PDB código 1A67) y empleando el servidor pyDockWeb [54]; el grupo tiometil de Cys<sup>25</sup> y Cys<sup>117</sup> de la estructura de la enzima fue previamente removido. Se empleó la versión restrictiva de pyDock (RST), la cual permite imponer restricciones sobre posibles contactos interfaciales.

Para llevar a cabo el modelado por reconocimiento molecular (*docking*) del heterodímero quimopapaína–CEW, el modelo atómico del complejo homólogo papaína–estefina B (PDB código 1STF) [22] se empleó como plantilla para establecer las restricciones requeridas en el servidor pyDock (RST). En base a un alineamiento de la secuencia de aminoácidos respectivamente entre las enzimas (quimopapaína y papaína) y los inhibidores (cistatina de pollo y estefina B) se establecieron como posibles residuos en contacto los siguientes: Para quimopapaína N<sup>18</sup>, E<sup>19</sup>, G<sup>20</sup>, A<sup>21</sup>, C<sup>22</sup>, G<sup>23</sup>, C<sup>63</sup>, K<sup>64</sup>, G<sup>65</sup>, G<sup>66</sup>, Y<sup>67</sup> y Q<sup>142</sup>; Para cistatina de pollo: G<sup>9</sup>, A<sup>10</sup>, P<sup>11</sup>, Q<sup>53</sup>, L<sup>54</sup>, V<sup>55</sup>, S<sup>56</sup> y G<sup>57</sup>.

Los modelos obtenidos con mayor puntuación fueron superpuestos a la estructura cristalográfica del complejo papaína–estefina B, y se procedió a calcular su desviación del valor cuadrático medio (RMSD) respecto a la posición de los carbonos alfa (C<sub>α</sub>). El modelo con el más bajo RMSD fue seleccionado como estructura de inicio para continuar con un refinamiento con simulación de dinámica molecular.

Las simulaciones con dinámica molecular (MD, por las siglas de *molecular dynamics*), fueron realizadas empleando el software GROMACS 4.5.4 [55] implementando el campo de fuerza OPLS [56]. El modelo quimopapaína–CEW fue colocado en el centro de una caja dodecaédrica periódica, con 10 Å entre la estructura del complejo y el borde de la caja. Se requirieron un total de 24899 moléculas de agua TIP3P para llenar la caja, y 12 iones cloruro para neutralizar la carga neta de la proteína. Previo a las simulaciones de MD, el sistema fue relajado por minimización de energía, seguida por 100 ps de equilibrio térmico restringiendo la posición de los átomos pesados de proteína con una constante de fuerza armónica de 1000 KJ mol<sup>-1</sup> nm<sup>-1</sup>. La simulación MD fue realizada empleando un conjunto NPT a 300 K y 1.0 bar

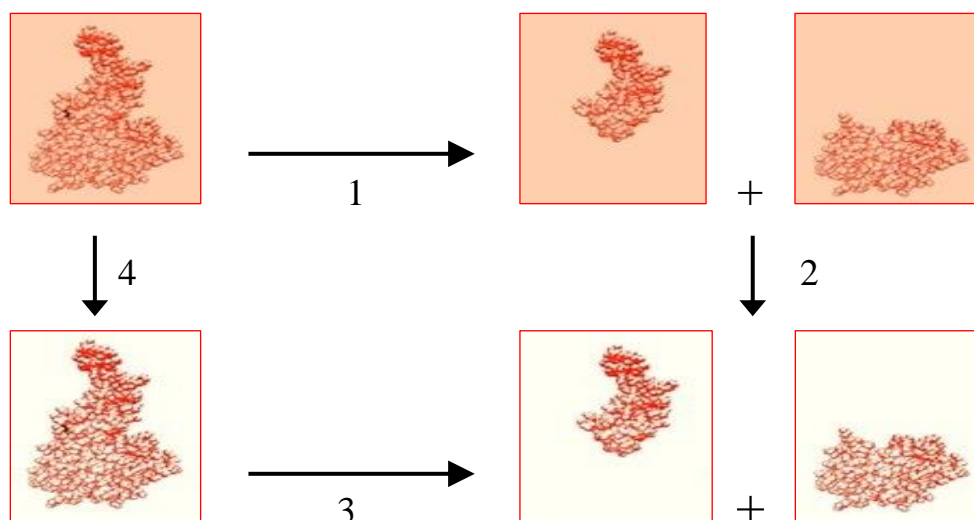
por 100 ns. El algoritmo LINCS fue aplicado para restringir la longitud de todos los enlaces covalentes [57], empleando un paso de tiempo de 2 fs. Para las interacciones de corto alcance y de van der Waals se aplicó un radio de corte de 1.2 nm, mientras que las interacciones electrostáticas de largo alcance fueron tratadas con el método de PME (Particle Mesch Ewald) [58].

Una vez obtenido el modelo del complejo quimopapaína–cistatina, se procedió a identificar en la interfase los residuos ionizables (Asp, Glu, Tyr, Cys, His, Lys y Arg) a un radio de corte  $\leq 5.0 \text{ \AA}$  entre el residuo y la cadena opuesta. Los estados de ionización de los grupos protonables presentes en la cadena de proteína libre y asociada se establecieron utilizando el valor de  $pK_a$  estimado por PROPKA [59]. El programa NACCESS fue usado con un radio de prueba de  $1.4 \text{ \AA}$  para calcular el área de superficie accesible al solvente para cada residuo en las proteínas separadas (%  $SSA^S$ ) y asociada (%  $SSA^C$ ) [60].

### **3.4.2. Cálculo de la energía electrostática de unión para el complejo quimopapaína–CEW.**

El cálculo de la energía de unión electrostática ( $\Delta G_{b,elec}$ ) del complejo bimolecular se realizó asumiendo una conformación rígida [33]. En el diagrama termodinámico de unión presentado en la Figura 3.7, se ilustra la asociación en términos de transferencia de energía libres a partir de un entorno dieléctrico homogéneo (donde las interacciones son descritas por la ley de Coulomb) a un entorno dieléctrico heterogéneo con diferentes constantes dieléctricas en la proteína y en el solvente [61].

La energía de unión electrostática ( $\Delta G_{b,elec}$ ) es obtenida a partir de la suma de dos componentes: la energía de solvatación ( $\Delta \Delta G_{solv}$ ) y la energía Coulombica ( $\Delta \Delta G_{coul}$ ). Ambas energías para la unión de quimopapaína y CEW fueron determinadas usando el programa Adaptive Poisson-Boltzmann Solver (APBS) [61], como se describe previamente por Lira-De León *et al.* (2013) [62]. Los estados de protonación de los residuos ionizables se asignaron con PROPKA [59] a los diferentes valores de pH en estudio, empleando el campo de fuerza PARSE [63] para asignar el radio iónico y las cargas atómicas. Los parámetros de la proteína fueron obtenidos con el servidor PDB2PQR [64], empleando una constante dieléctrica de 78.0 y 4.0 para el agua y la proteína, respectivamente.



**Figura 3.7** Diagrama termodinámico que ilustra el procedimiento estándar para calcular la contribución electrostática a la energía libre de unión de dos moléculas con cuerpo rígido.

A) La contribución de la energía de solvatación ( $\Delta\Delta G_{solv}$ ) involucra los pasos 4 y 2 definida por:

$$\Delta\Delta G_{solv} = \Delta G_4 - \Delta G_2 = \Delta G_{solv}^{complejo} - (\Delta G_{solv}^{mol\ 1} + \Delta G_{solv}^{mol\ 2}).$$

B) La contribución de energía Coulombica ( $\Delta\Delta G_{coul}$ ) está dado por 1:

$$\Delta\Delta G_{coul} = -\Delta G_1 = \Delta G_{coul}^{complejo} - (\Delta G_{coul}^{mol\ 1} + \Delta G_{coul}^{mol\ 2}).$$

C) La energía libre de unión esta descrita por el paso 3:

$$\Delta G_{b,elec} = -\Delta G_3 = \Delta G_4 - (\Delta G_1 + \Delta G_2).$$

### 3.4.3. Comparación de la interfase del modelo quimopapaína–CEW con estructuras homólogas cisteinproteasa–cistatina.

La composición de la interfase del modelo quimopapaína cistatina de pollo fue comparada con seis estructuras homólogas cisteinproteasa–cistatina: papaína–CEW, papaína–estefina B<sup>a</sup>, catepsina B–estefina A<sup>b</sup>, catepsina H–estefina A<sup>c</sup>, catepsina L–estefina A<sup>d</sup>, catepsina V–estefina A<sup>e</sup>. Se construyó la estructura del heterodímero papaína–cistatina de pollo por modelado de homología empleando el programa MOE, utilizando como plantilla la estructura del modelo Quimopapaína–CEW. Para los otros complejos, se usaron las estructuras cristalográficas obtenidas de Protein Data Bank con los códigos de identificación: <sup>a</sup>1STF, <sup>b</sup>3K9M, <sup>c</sup>1NB5, <sup>d</sup>3KSE, <sup>e</sup>3KFQ. Se obtuvieron los archivos PDB y borraron los heteroátomos (iones y moléculas de agua) y cadenas adyacentes contenidas en los archivos cristalográficos. Posteriormente, estas estructuras se sometieron a una minimización de energía con el campo de fuerzas YASARA, enviando las estructuras al servidor YASARA Energy Minimization [65].

En las estructuras en comparación se identificaron los residuos ionizables localizados en la interfase tomando como criterio un radio de corte de 5.0 Å de cualquier átomo de la cadena opuesta en la interfase, esta exploración de la interface se realizó con el programa VMD (Visual Molecular Dynamics). Las propiedades estructurales de la interfase a comparar fueron la presencia de puentes de hidrógeno a través de la interfase ( $N_{HB}$ ), pares iónicos a través de la interfase ( $N_{IP}$ ) y Área de superficie de contacto, calculada como el cambio en el área de superficie accesible al solvente entre el heterodímero y la suma de las proteínas libres usando el programa NACCESS [60].

### 3.5 Modelo modificado de transferencia de protón vinculado a la asociación: *MPLM* (Crnogorac *et al.*, 2001).

El fenómeno de la dependencia del pH en la asociación de quimopapaína con CEW se estudió con el modelo modificado de unión vinculado a la transferencia de protón (Crnogorac *et al.*, 2001) [34], designado en el presente trabajo como modelo *MPLM* (por sus siglas en inglés, Modified Proton-Linkage Model). El procedimiento consiste de tres partes.

- A) El primero consiste en determinar la dependencia de la constante de asociación ( $K_b$ ) con respecto a la variación de pH, lo cual se consiguió con la determinación de la  $K_b$  a diferentes valores de pH (ver, Sección 3.3.2).
- B) En la segunda parte, a partir de un modelo molecular construido con MD (ver, Sección 3.4.1) se identificaron los residuos ionizables en la interfase del heterodímero proteína–inhibidor y se estimaron sus valores de  $pK_a$ , tanto en la proteína en la forma asociada ( $pK_a^c$ ) como en las proteínas libres ( $pK_a^s$ ) utilizando el programa PROPKA [59]. Los residuos que presentaron un comportamiento similar en sus valores de  $pK_a$  en la forma libre y asociada se agruparon en grupos isoacídicos (*composite pK<sub>a</sub>*) [34].
- C) En la tercera etapa, se procedió a evaluar los valores de  $pK_a^c$  y  $pK_a^s$  (predichos por PROPKA) en la ecuación del modelo *MPLM* (Ecuación 3.3), para obtener curvas que describieran el perfil de los datos experimentales de  $K_b$  con respecto al cambio de pH. Se realizaron simulaciones de las curvas resultantes de las combinaciones de tres grupos de residuos isoacídicos, en el intervalo de pH de 2 a 12, con el software Wolfram Mathematica 9. Una vez identificada la posible combinación de grupos

isoacídicos, se procede a realizar el ajuste de los datos experimentales de  $K_b$  obtenidos a los diversos valores de pH al modelo *MPLM* con el software Origin Pro 8.6.

---

**Ecuación 3.3** Ecuación del modelo *MPLM* (detalles de la derivación ver apéndice D).

---

$$K_b = K_0 \frac{\sum_{i=0}^M \sum_{j=1}^{p_i} [10^{-\text{pH}}]^{-a_i^j} 10^{-\sum_{k=1}^M x_{ik}^j m_k \text{p}K_{a_k}^c}}{\sum_{i=0}^M \sum_{j=1}^{p_i} [10^{-\text{pH}}]^{-a_i^j} 10^{-\sum_{k=1}^M x_{ik}^j m_k \text{p}K_{a_k}^s}}$$

Donde:  $K_b$  es la constante de unión;  $K_0$ , es la constante de unión en el estado completamente protonado;  $M$ , es el número de grupos isoacídicos indexados por  $k$ ;  $\text{p}K_a$ , constante de disociación ácida de los residuos en la interfase del heterodímero:  $\text{p}K_a^s$  para la proteína en la forma separada o libre y  $\text{p}K_a^c$  para la forma combinada o asociada; El índice  $i$  ( $i \leq M$ ) y  $j$  representan el número de residuos isoacídicos en el estado desprotonado y protonado, respectivamente;  $a_i^j$  es el número de protones disociados a partir de la proteína;  $x_{ik}^j$  puede tomar el valor de 1 o 0 dependiendo si el sitio  $i$  es protonado o no;  $m$  es el número de residuos que constituyen el grupo isoacídico.

---

En el proceso de asociación de las proteínas en estudio puede estar presente la liberación o captación de  $H^+$ . El número promedio de  $H^+$  ( $q$ ) está dado por la ecuación 3.4, valores positivos y negativos de  $q$  corresponden a liberación y captación, respectivamente, de iones  $H^+$  bajo la asociación de las proteínas en estudio.

---

**Ecuación 3.4** Número de iones  $H^+$  ( $q$ ) transferidos a partir del sistema de estudio hacia el amortiguador de pH.

---

$$\frac{d \log K}{d \text{pH}} = q$$

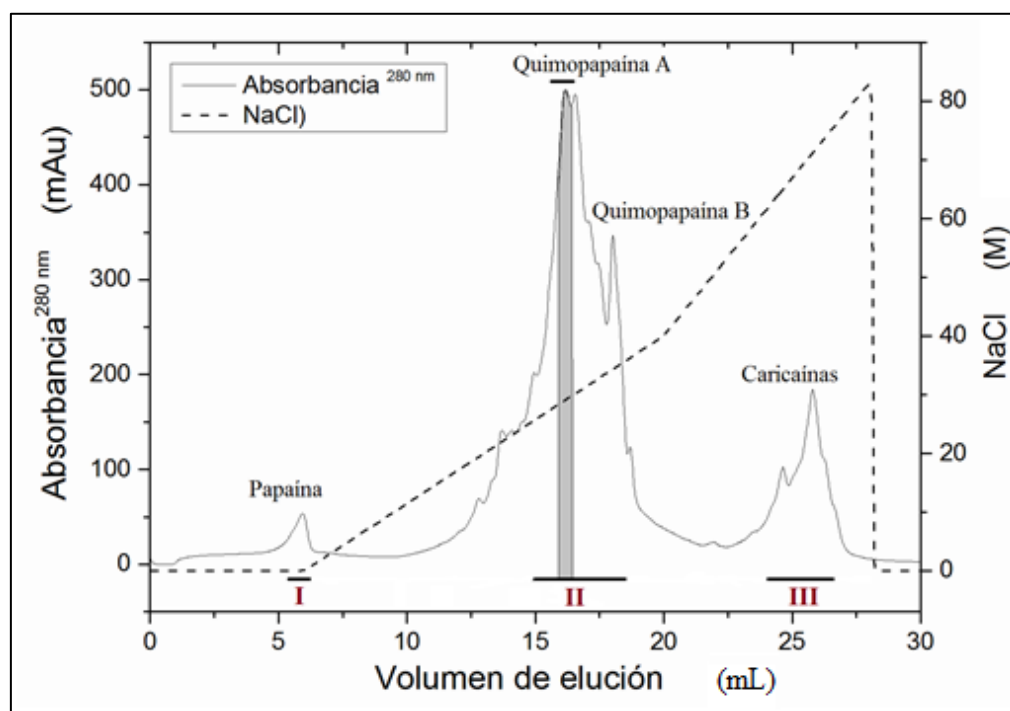
Donde:  $q$  es el número de  $H^+$  liberados o captados bajo la asociación;  $K$ , constante de unión calculada experimentalmente; pH, logaritmo negativo de la concentración de protones.

---

## 4. RESULTADOS.

### 4.1. Purificación de SCM-quimopapaína.

Mediante cromatografía de intercambio catiónico es muy fácil identificar la SCM-quimopapaína. En un cromatograma típico de la elución de proteínas contenidas en la preparación de quimopapaína (Figura 4.1) se resuelven tres zonas de elución. La zona I corresponde a la elución de papaína y proteínas no unidas a la columna, la zona II a quimopapainas y la zona III a caricainas. La proteína de interés eluye en la zona de quimopapainas, en esta región se han identificado dos picos correspondientes a quimopapaína A y quimopapaína B [11,44].

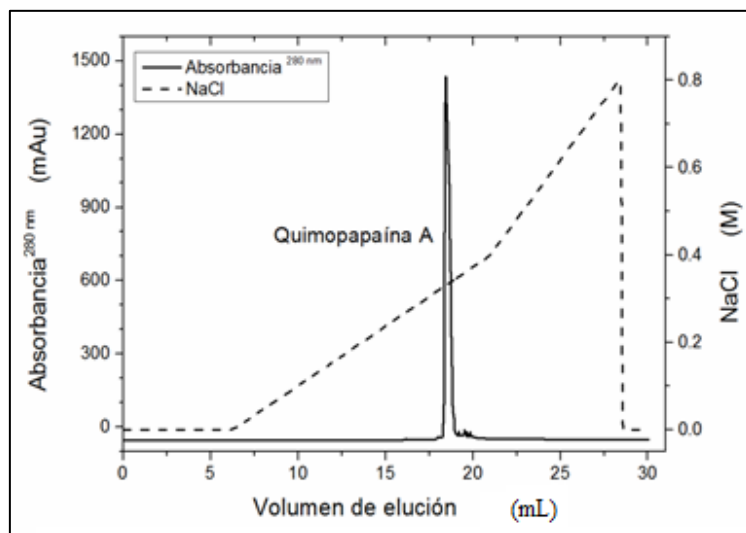


**Figura 4.1** Cromatograma de extracto de quimopapaína. Cromatografía de intercambio catiónico: columna Mono S 5/50 GL, amortiguador de fosfato de sodio 50 mM, pH 7.0. El área sombreada en color gris indica la fracción de quimopapaína colectada

La quimopapaína es una proteína que existe en múltiples formas (natural y otras que sufren modificaciones durante la preparación comercial de látex) [11], la forma A es la proteína de interés (ver Sección 3.1). En el cromatograma presentado en la Figura 4.1, se indica la fracción colectada de quimopapaína (área en tono gris) para realizar los experimentos con espectroscopia UV-Visible, DC y fluorescencia. El rendimiento de purificación de SCM-quimopapaína a partir del extracto comercial fue del 10 %.

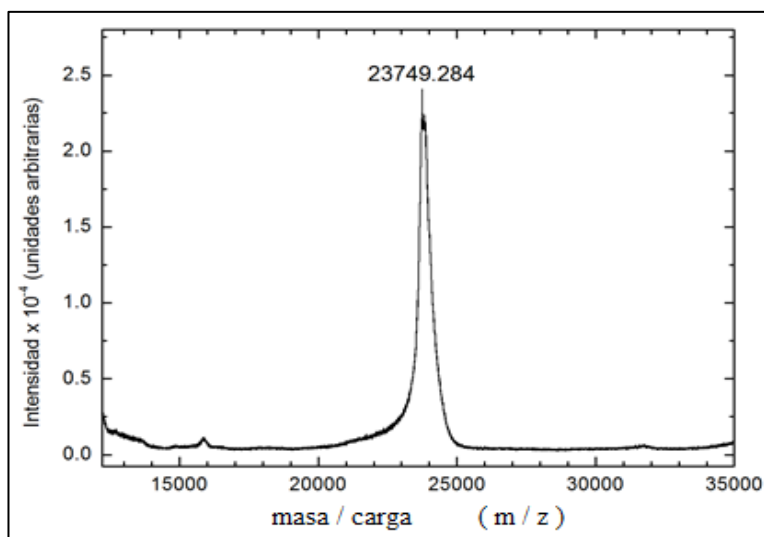
#### 4.1.1. Pruebas de pureza de SCM-quimopapaína.

En el cromatograma presentado en la Figura 4.2, se aprecia la elución de un pico bien definido, resultado de la cromatografía de la fracción de SCM-quimopapaína A.



**Figura 4.2** Cromatograma de SCM-quimopapaína. Cromatografía de intercambio catiónico: columna Mono S 5/50 GL, amortiguador de fosfato de sodio 50 mM, pH 7.0.

El resultado de la espectrometría de masas de la SCM-quimopapaína, ver Figura 4.3, muestra una molécula con una masa de 23749.3 Da que corresponde a la masa molecular de SCM-quimopapaína (23752 Da).

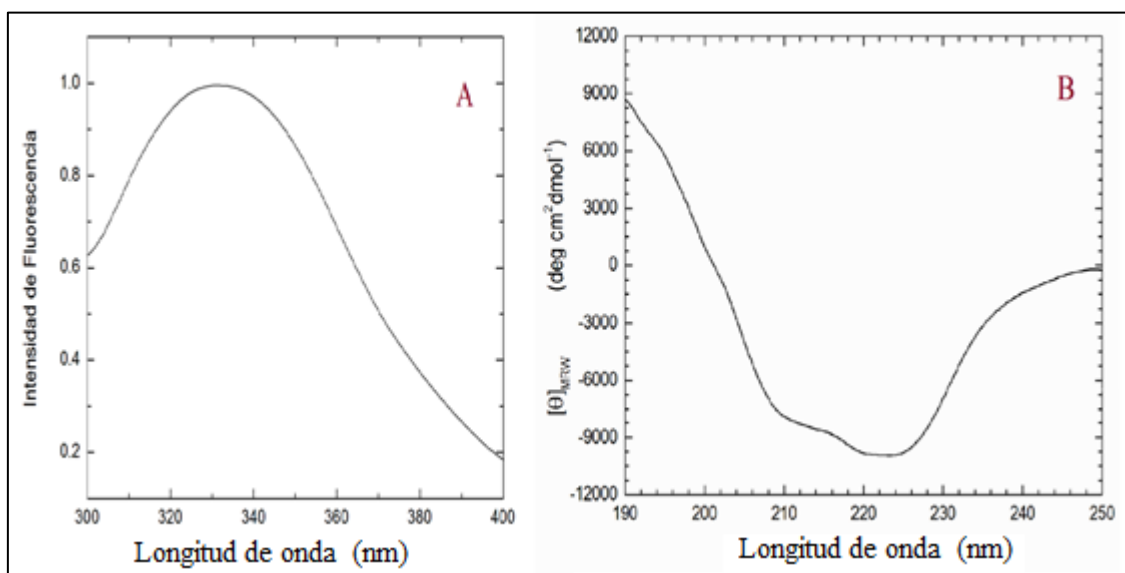


**Figura 4.3** Espectro de masas MALDI-TOF de la SCM-quimopapaína (usando HCCA como matriz). El espectro se obtuvo en modo lineal positivo y no fue resuelto el patrón isotópico.

#### 4.1.2. Propiedades espectrales de SCM-quimopapaína purificada.

La quimopapaína contiene cuatro triptófanos ( $W^7$ ,  $W^{181}$  y  $W^{185}$ , localizados en el dominio R de la molécula y el  $W^{26}$  en el dominio L, ver Figura 1.3) y 15 tirosinas. En la Figura 4.4-A se presenta un espectro típico de emisión de fluorescencia de SCM-quimopapaína, el cual exhibe un máximo alrededor de 330 nm.

El espectro de DC en la región UV-lejano de SCM-quimopapaína (Figura 4.4-B) exhibe un mínimo a una longitud de onda de 208 nm ( $\sim -6800 \text{ deg cm}^2 \text{ dmol}^{-1}$ ) y uno más prominente alrededor de 222 nm ( $\sim -9900 \text{ deg cm}^2 \text{ dmol}^{-1}$ ), indicio de la presencia de  $\alpha$ -hélice. Estos resultados coinciden con los reportados por Solís-Mendiola *et al.* (1989,1992) [44,66].

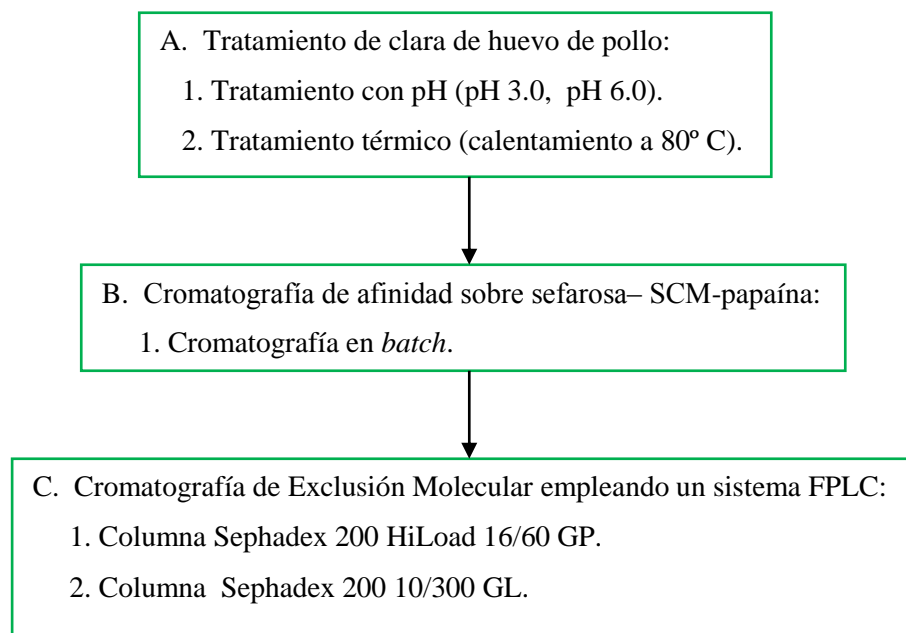


**Figura 4.4** Propiedades espectrales de SCM-quimopapaína. (A) Espectro de emisión de fluorescencia (por ejemplo, concentración de enzima 0.1  $\mu\text{M}$ , amortiguador de pH 7.4, Tris-base 50 mM, NaCl 0.1 M y 25 °C). (B) Espectro de DC de SCM-quimopapaína en la región UV-lejano (por ejemplo, concentración de enzima 0.1 mg/mL, amortiguador de pH 7.5, fosfato de sodio 50 mM y 25 °C).



## 4.2. Purificación de CEW.

Se logró purificar cistatina a partir de clara de huevo de pollo con el método de purificación desarrollado en el presente trabajo. El método implementado consiste en tres etapas (ver Figura 4.5): a) Tratamiento de clara de huevo (Sección 3.2.1), b) Cromatografía de afinidad (Sección 3.2.3) y c) Cromatografía por exclusión molecular (Sección 3.2.4).



**Figura 4.5** Método de purificación de cistatina a partir de clara de huevo de pollo.

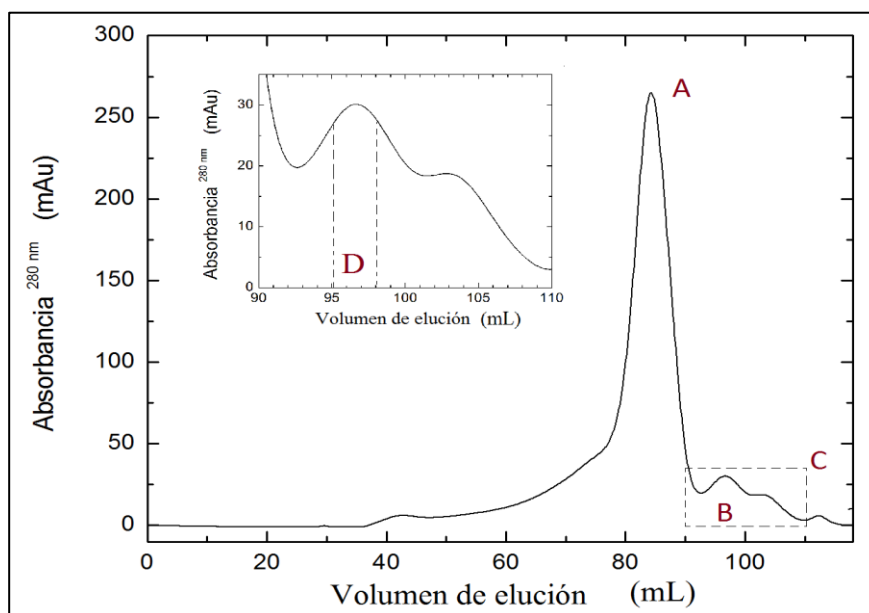
El rendimiento de purificación de CEW a partir de clara de huevo fue del 8.5 %. El valor fue obtenido por lote de clara de huevo procesado, cada lote consistió en procesar la cantidad de 30 piezas de huevo, de los cuales se obtuvieron aproximadamente 2.1 mg de cistatina pura.

En las etapas intermedias de purificación de cistatina no me fue posible cuantificar con precisión la cantidad de cistatina de interés y determinar su pureza. Por tal motivo, sólo se presenta el rendimiento al final de la purificación.

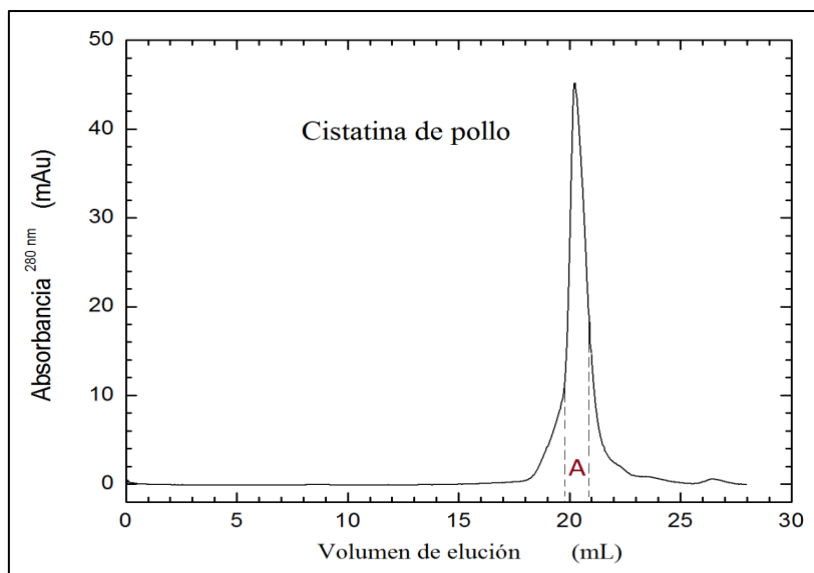
En el método convencional de purificación de cistatina, se obtienen rendimientos del orden del 20 %, la mayor pérdida de cistatina en las estepas de purificación ocurre durante la cromatografía de afinidad [51].

En la Figura 4.6 se presenta un cromatograma típico de la elución de la solución de cistatina concentrada obtenida de cromatografía de afinidad a través de una columna de exclusión molecular. Tanto el pico A como la zona de elución resaltada en el inserto presentan actividad inhibitoria de cisteinoproteasas (Prueba de actividad inhibitoria Sección 3.2.5.A). La fracción de interés se encuentra en la elución del pico B, la fracción colectada es señalada en el inserto como la fracción D.

La fracción D fue concentrada y cargada en una columna de exclusión molecular grado analítico. La fracción A indicada en la en Figura 4.7 (cromatograma típico) corresponde a la cistatina de pollo purificada.



**Figura 4.6** Cromatograma de exclusión molecular de la solución de cistatina obtenida de la cromatografía de afinidad. Columna de exclusión molecular Superdex 200 HiLoad 16/60 GP, amortiguador de elución de fosfato de sodio 30 mM, NaCl 50mM, pH 7.0; **A**, Indica el pico de elución de proteína con masa molecular aproximada de 40 KDa; **B**, Indica el pico de elución de proteína con masa molecular aproximada de 13 KDa; **C**, indica la zona donde se capturó el inserto; **D**, Fracción colectada de interés.



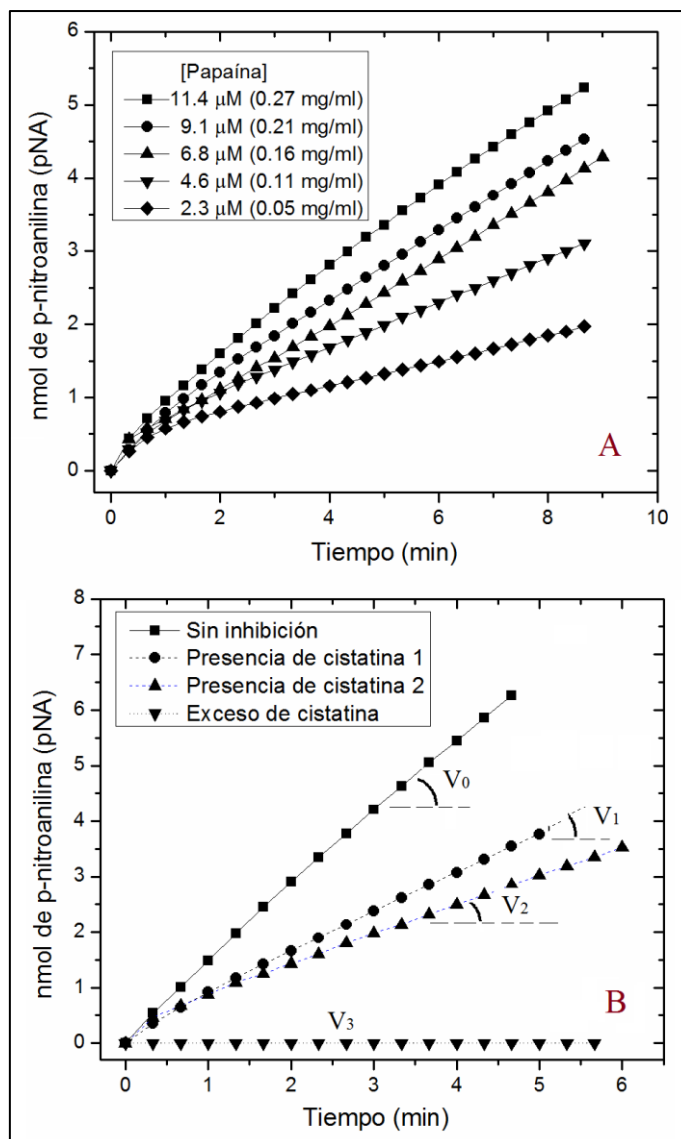
**Figura 4.7** Cromatograma de exclusión molecular de CEW. Columna de exclusión molecular Superdex 200 10/300 GL, amortiguador de elución de fosfato de sodio 30 mM, NaCl 50mM, pH 7.0; A, indica la fracción de interés, correspondiente a CEW.

#### 4.2.1. Ensayos de actividad inhibitoria y pruebas de pureza de CEW.

Se logró estandarizar el método colorimétrico BApNA [27] para verificar la actividad inhibitoria de cistatina. Las cinéticas de hidrólisis de sustrato BApNA por papaína presentan linealidad a concentraciones de 4.6 a 12  $\mu$ M de enzima a partir de 2.5 min en cinéticas de 10 minutos (Figura 4.8.A), en las condiciones del ensayo (Sección 3.2.5.A).

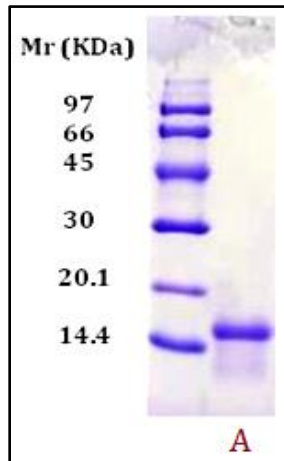
Los ensayos realizados en presencia y ausencia de CEW, sólo nos dieron información cualitativa de la presencia o ausencia de inhibición (Figura 4.8.B).

Como parámetro de comparación en estos experimentos de inhibición, se calculó el cambio en la velocidad de formación de producto ( $\Delta V$ ) como:  $\Delta V = V_i/V_0$ , donde,  $V_0$  y  $V_i$  son las velocidades de formación de producto en cinéticas de hidrólisis de BApNA en ausencia y presencia de CEW, respectivamente (ver ejemplo, Figura 4.8.B). Cuando  $\Delta V$  es igual a 1, no se presenta inhibición y cuando se obtiene un valor entre 1 y 0, se puede determinar el porcentaje de papaína inhibida. Estos experimentos de inhibición se realizaron con papaína comercial parcialmente purificada, por tal motivo, no se presentan cálculos de parámetros cinéticos de inhibición.



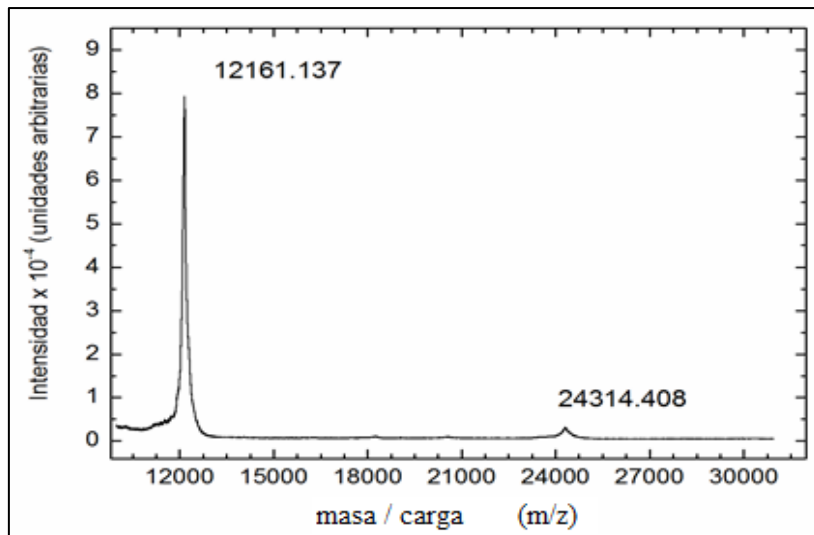
**Figura 4.8** Cinéticas típicas de la hidrólisis de sustrato BApNA por papaína. a) Cinéticas de hidrólisis de sustrato a diferentes concentraciones de papaína. b) Cinéticas de hidrólisis de papaína: V indica la velocidad de formación de producto (pNA), el subíndice 0 indica ausencia de cistatina, el subíndice 1, 2 presencia de cistatina y subíndice 3 un exceso molar de cistatina con respecto a la cantidad de papaína.

En la Figura 4.9 se presenta el gel resultado de la electroforesis (SDS-PAGE) de CEW después de ser pasada por la columna de exclusión molecular (Superdex 200 10/300 GL), se observa una banda de peso molecular aparente de 14.4 KDa correspondiente a la CEW. Este resultado ya se ha reportado por otros grupos de investigadores [27,29].



**Figura 4.9** Gel de electroforesis SDS-PAGE de CEW (12%, tinción con azul coomassie). Carril A, cistatina de pollo purificada.

El resultado de la espectrometría de masas de CEW (Figura 4.10) reveló la presencia de una molécula con masa de 12167.1 Da y una señal correspondiente a 24314.4 Da. La primera corresponde a la masa molecular de cistatina de pollo cortada en el extremo amino terminal ( $G^9$ ), determinado a partir de la secuencia de estructura primaria (resultado ya reportado por otros grupos de investigadores [19]). La segunda señal corresponde a una pequeña fracción de cistatina que se encuentra en forma dimérica, no hay información reportada al respecto.

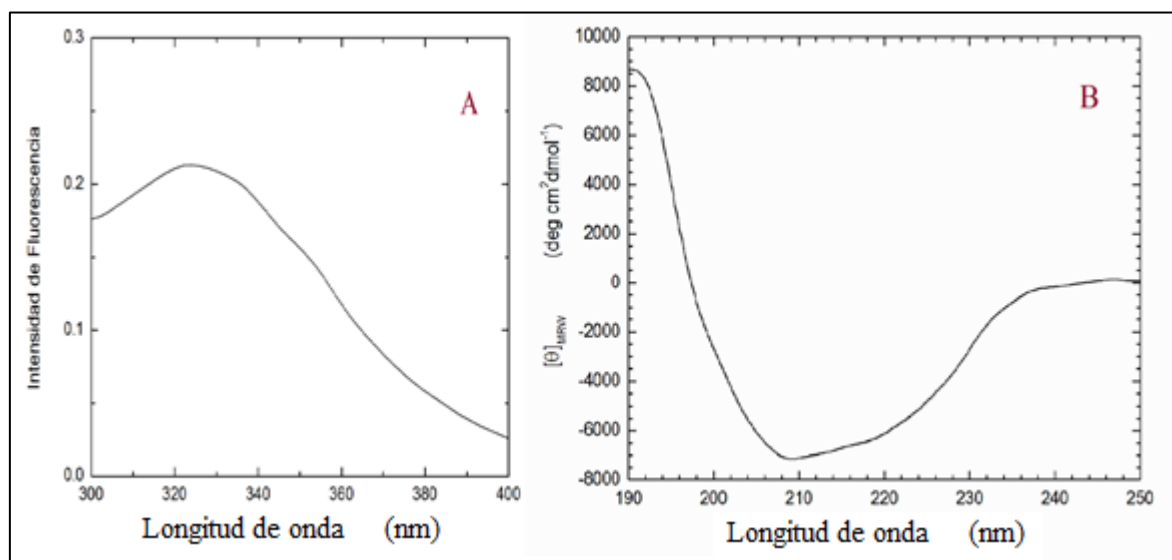


**Figura 4.10** Espectro de masas MALDI-TOF de CEW (usando HCCA como matriz). El espectro se obtuvo en modo lineal positivo y no fue resuelto el patrón isotópico.

### 4.2.2. Propiedades espectrales de CEW purificada.

El espectro típico de emisión de fluorescencia de CEW es presentado en la Figura 4.11.A, exhibe un máximo cercano a 325 nm. La molécula de CEW contiene en su estructura un triptófano ( $W^{104}$ , ver su localización en Figura 1.5) y cinco tirosinas.

El espectro de DC en la región UV-lejano de CEW (Figura 4.11.B) exhibe un mínimo a una longitud de onda de 208 nm ( $\sim -7000 \text{ deg cm}^2 \text{ dmol}^{-1}$ ) y un ligero hombro alrededor de 222 nm ( $\sim -5600 \text{ deg cm}^2 \text{ dmol}^{-1}$ ), elementos que indican que la molécula está compuesta de estructura  $\alpha$ -hélice y/o hojas  $\beta$ . El espectro tiene una forma similar a la reportada por Lindahl *et al.* (1988) [30].



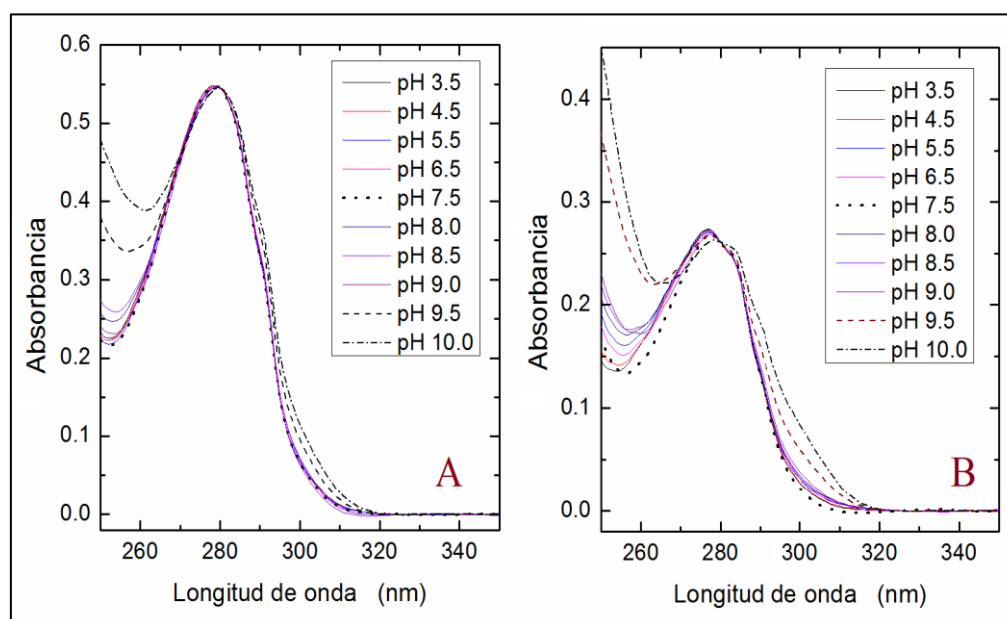
**Figura 4.11** Propiedades espectrales de CEW. (A) Espectro de emisión de fluorescencia (por ejemplo, concentración de cistatina 0.1  $\mu\text{M}$ , amortiguador de pH 7.4, Tris-base 50 mM, NaCl 0.1 M y 25  $^{\circ}\text{C}$ ). (B) Espectro de DC de CEW en la región UV-lejano (por ejemplo, concentración de cistatina 0.1 mg/mL, amortiguador de pH 7.5, fosfato de sodio 50 mM y 25  $^{\circ}\text{C}$ ).

### 4.3. Cambios espectroscópicos que acompañan la interacción de CEW y SCM-quimopapaína a diferentes valores de pH, a 25 °C.

#### 4.3.1. Cambios espectrales en absorbancia en la región UV-visible y DC en la región UV-lejano a pH de 3.5 a 10.0.

Los espectros de absorbancia en la región ultravioleta-visible cercano de SCM-quimopapaína (Figura 4.12.A) y CEW (Figura 4.12.B), a valores de pH de 3.5 a 9.0 presentan un espectro típico de proteína, ambos exhiben un máximo en 280 nm. La exploración de los espectros (en ambas proteínas) en el visible cercano no muestra evidencia de la presencia de grupos cromóforos prostéticos (grupos de naturaleza no proteica) y no se presenta señal de absorbancia por arriba de 320 nm [47,67].

Los espectros de SCM-quimopapaína (Figura 4.12.A) y CEW (Figura 4.12.B) a valores de pH de 9.5 y 10.0 respectivamente, exhibe un ligero ensanchamiento notorio con respecto al espectro a pH 7.5, este efecto es resultado del cambio en el entorno de los grupos cromóforos (Tyr y Trp) presentes en la respectiva estructura de SCM-quimopapaína y CEW. La completa ionización de tirosina causa el típico corrimiento batocrómico (corrimiento del máximo de absorbancia a 290 nm) [67], ausente en los espectros obtenidos.



**Figura 4.12** Efecto del pH sobre las propiedades espectrales de SCM-quimopapaína y CEW: Espectros de absorbancia en la región UV cercano (250-350 nm, 0.3 mg/mL y 25°C) de SCM-quimopapaína (A) y CEW (B) y a diferentes valores de pH (ver Sección 3.3.1).

Analizando la razón de la medición a 280 nm ( $A^{280}$ ) con respecto a 260 nm,  $A^{280/260}$  y a 292 nm,  $A^{280/292}$  a valores entre pH 3.5 y 9.0 (valores de la Tabla 4.1), se obtiene para quimopapaína:  $A^{280/260} = 1.89 \pm 0.047$  y para  $A^{280/292} = 2.02 \pm 0.022$ . Estos valores son comparables con los reportados para pH de 2.1 a 7.38 [67], para los cuales se obtiene la razón de  $1.95 \pm 0.056$  y  $2.08 \pm 0.114$ . Cuando ocurre el fenómeno batocrómico originado por la ionización de tirosina para quimopapaína pH 13.0 [67] se registra a  $Abs^{280/260} = 0.700$  y para  $Abs^{280/292} = 0.936$ . Con respecto a cistatina no encontré en la bibliografía consultada parámetros para comprar los resultados obtenidos.

Con respecto a las mediciones de absorbancia a valores de pH 9.5 y 10.0 se presentan ligeras señales de que el entorno de los grupos cromóforos está cambiando, posiblemente a la ionización de Tyr. Sin embargo, no se aprecia un cambio en la medición a 280 nm (punto isobéptico) para ambas proteínas (ver Figura 4.12). Por lo tanto, para ambas proteínas no es necesario corregir el coeficiente de extinción molar para determinar la concentración de proteína a los diferentes valores de pH.

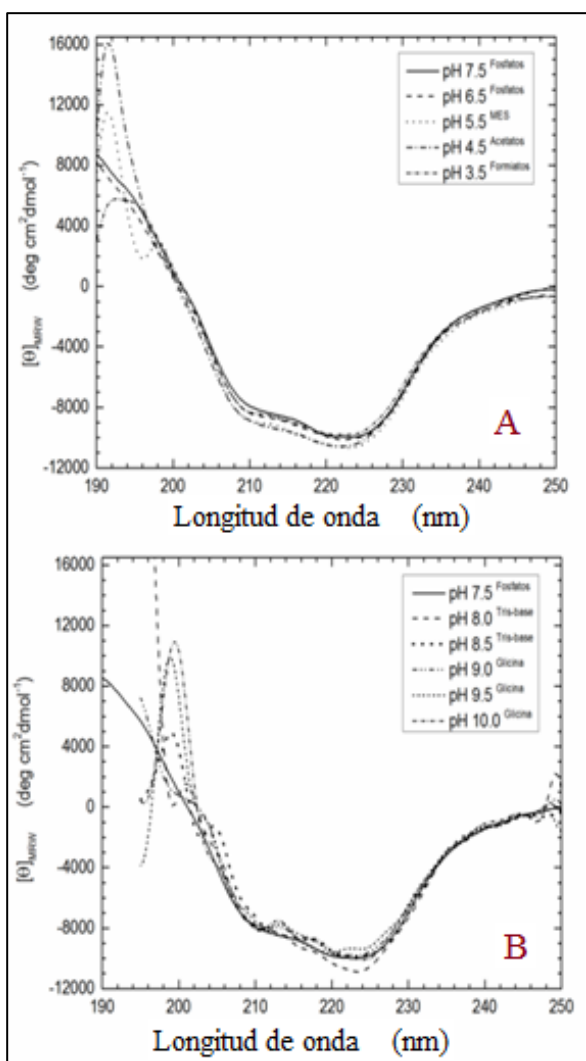
**Tabla 4.1** Razón de la medición de absorbancia en UV de proteínas ( $Abs^{280/260}$  y  $Abs^{280/292}$ ).

pH	SCM-quimopapaína		CEW	
	$Abs^{280/260}$	$Abs^{280/292}$	$Abs^{280/260}$	$Abs^{280/292}$
3.5	1.874	2.031	1.611	2.559
4.5	1.883	2.032	1.606	2.528
5.5	1.926	2.030	1.515	2.479
6.5	1.901	2.026	1.578	2.447
7.5	1.956	2.026	1.805	2.586
8.0	1.848	1.991	1.470	2.440
8.5	1.823	1.974	1.456	2.423
9.0	1.945	2.020	1.514	2.320
9.5	1.595	1.795	1.129	1.953
10.0	1.399	1.693	1.041	1.636
(3.5 - 9.0) <sup>1</sup>	$1.89 \pm 0.047$	$2.02 \pm 0.022$	$1.57 \pm 0.112$	$2.47 \pm 0.085$

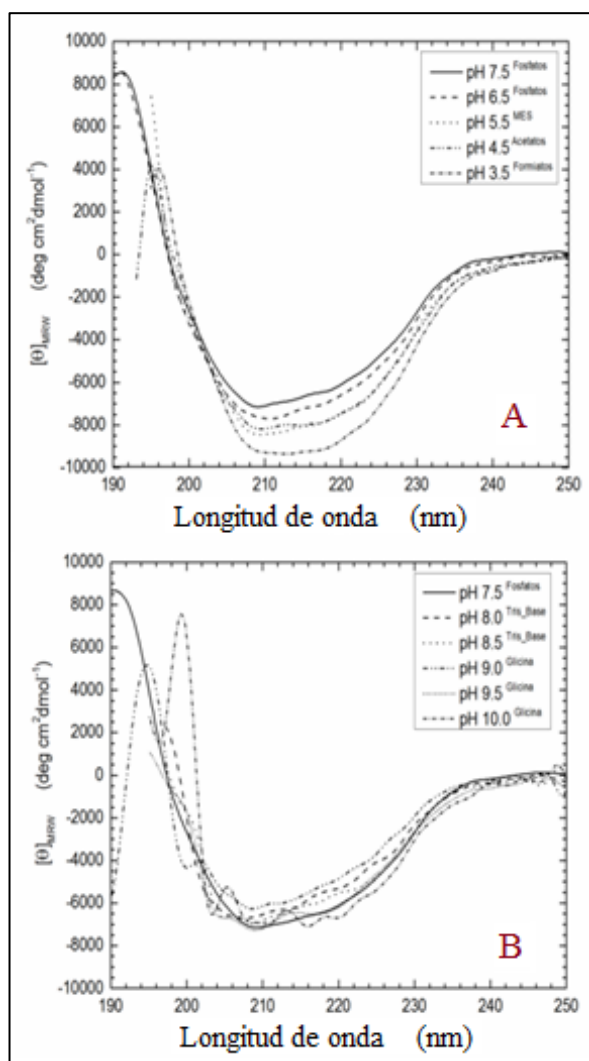
Donde: <sup>1</sup> Intervalo de pH en que se efectuó el análisis estadístico; la incertidumbre representa la desviación estándar de las mediciones; Abs, absorbancia, el superíndice indica la longitud de onda de las mediciones .



Los espectros de DC en la región UV-lejano o amida (170–250 nm) correspondientes a quimopapaína y CEW se presentan en la Figura 4.13 y Figura 4.14, respectivamente. Los espectros de DC de quimopapaína (Figura 4.13) no exhiben diferencias considerables con respecto al espectro obtenido a pH 7.5, tanto en la magnitud de la señal como en la posición de los dos mínimos (208 nm y 222 nm). En el caso de los espectros de DC de CEW (Figura 4.14) presentan pequeñas diferencias en magnitud, la localización del mínimo a 208 nm (identificado a pH 7.5) no presenta corrimiento a otra longitud de onda en los diversos valores de pH.



**Figura 4.13** Espectros de CD de SCM-quimopapaína en la región UV lejano (190-250 nm) a diferentes valores de pH (ver Sección 3.3.1).



**Figura 4.14** Espectros de CD de CEW en la región UV lejano (190-250 nm) a diferentes valores de pH (ver Sección 3.3.1).

Los espectros correspondientes a quimopapaína y CEW, presentados en la Figura 4.13 y 4.14, fueron analizados con el software DicroProt 2000, (empleando el método CONTIN [68]), los resultados son presentados en la Tabla 4.2. No se detectaron cambios relevantes en la composición de estructura secundaria en ambas proteínas a los diferentes valores de pH analizados. Los mayores cambios se presentan en CEW a pH de 10.0, en base al análisis hay un mayor ordenamiento de estructura hoja  $\beta$ .

**Tabla 4.2** Estructura secundaria de SCM-quimopapaína y CEW a diferentes valores de pH (análisis efectuado con el método CONTIN).

pH	Hélice (%)	Hoja $\beta$ (%)	Remanente (%)	Factor	$[\theta]_{208} / [\theta]_{222}$
SCM-quimopapaína <sup>a</sup>					
<b>3.5</b>	24 $\pm$ 2.3	24 $\pm$ 2.9	52 $\pm$ 2.5	1.007	0.759
<b>4.5</b>	25 $\pm$ 2.4	22 $\pm$ 3.0	52 $\pm$ 2.6	1.010	0.760
<b>5.5</b>	28 $\pm$ 3.9	14 $\pm$ 6.4	57 $\pm$ 4.7	1.003	0.756
<b>6.5</b>	24 $\pm$ 2.4	23 $\pm$ 3.0	53 $\pm$ 2.6	1.008	0.737
<b>7.5</b>	23 $\pm$ 2.4	24 $\pm$ 3.0	53 $\pm$ 2.6	1.007	0.696
<b>8.0</b>	27 $\pm$ 4.1	15 $\pm$ 6.9	59 $\pm$ 5.0	1.003	0.649
<b>8.5</b>	23 $\pm$ 2.4	24 $\pm$ 3.0	53 $\pm$ 2.6	1.004	0.607
<b>9.0</b>	24 $\pm$ 2.3	24 $\pm$ 3.0	52 $\pm$ 2.6	1.007	0.712
<b>9.5</b>	22 $\pm$ 3.7	24 $\pm$ 6.5	54 $\pm$ 4.7	0.999	0.697
<b>10</b>	22 $\pm$ 3.9	27 $\pm$ 6.7	51 $\pm$ 4.9	0.999	0.666
CEW <sup>b</sup>					
<b>3.5</b>	21 $\pm$ 1.8	27 $\pm$ 2.7	52 $\pm$ 2.3	1.014	1.102
<b>4.5</b>	17 $\pm$ 1.5	31 $\pm$ 2.5	52 $\pm$ 2.1	1.007	1.133
<b>5.5</b>	18 $\pm$ 1.5	30 $\pm$ 2.5	51 $\pm$ 2.1	1.010	1.163
<b>6.5</b>	16 $\pm$ 1.3	33 $\pm$ 2.3	51 $\pm$ 2.0	1.001	1.214
<b>7.5</b>	15 $\pm$ 0.8	35 $\pm$ 1.2	50 $\pm$ 1.4	0.952	1.244
<b>8.0</b>	10 $\pm$ 1.9	38 $\pm$ 4.0	51 $\pm$ 3.0	0.997	1.354
<b>8.5</b>	11 $\pm$ 1.3	36 $\pm$ 2.3	52 $\pm$ 2.0	0.987	1.275
<b>9.0</b>	11 $\pm$ 1.0	39 $\pm$ 2.0	50 $\pm$ 1.8	0.970	1.424
<b>9.5</b>	13 $\pm$ 1.2	36 $\pm$ 2.2	51 $\pm$ 1.9	0.987	1.275
<b>10</b>	19 $\pm$ 4.1	52 $\pm$ 8.0	29 $\pm$ 7.9	0.995	1.122

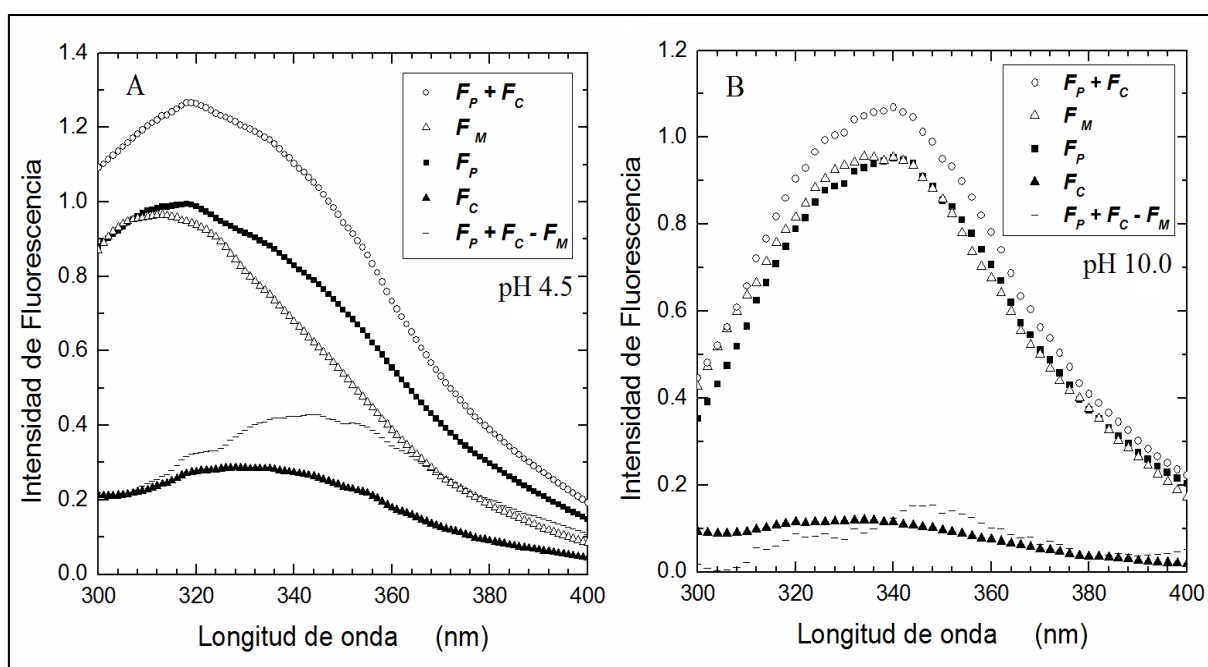
<sup>a</sup>Análisis de espectros de DC presentados en Figura 4.13;

<sup>b</sup>Análisis de espectros de DC presentados en Figura 4.14.

### 4.3.2. Cambios espectrales en fluorescencia que acompañan la interacción de SCM-quimopapaína y CEW: Constante de unión ( $K_b$ ) como función de pH.

Los espectros de emisión de fluorescencia fueron registrados de 300 a 400 nm. Todas las mediciones de fluorescencia se realizaron con a una longitud de onda de excitación de 280 nm, con la cual se excita a los fluoróforos intrínsecos tirosina y triptófano. La quimopapaína contiene en su estructura cuatro Trp y quince Tyr; por su parte, la CEW contiene un Trp y cinco Tyr.

En la Figura 4.15 se presenta un ejemplo de espectros de emisión de fluorescencia en condiciones de pH ácido (Figura 4.15-A) y en condiciones de pH alcalino (Figura 4.15-B). La posición del máximo de emisión de fluorescencia en los distintos amortiguadores y sus diversos valores de pH (ver Sección 3.3.2) dependerá de la polaridad del entorno y la movilidad del fluoróforo [69,70].



**Figura 4.15** Fluorescencia intrínseca de SCM-quimopapaína y CEW. Por ejemplo, en presencia de amortiguador de pH 4.5 con fuerza iónica 0.01M (A) y amortiguador glicina 50 mM, NaCl 0.1 M, pH 10.0 (B). Donde:  $F_P$ , fluorescencia de SCM-quimopapaína (0.1  $\mu\text{M}$ );  $F_C$ , fluorescencia de CEW (0.12  $\mu\text{M}$ );  $F_M$ , fluorescencia de la mezcla de SCM-quimopapaína (0.1  $\mu\text{M}$ ) y CEW (0.12  $\mu\text{M}$ );  $F_P + F_C$ , la suma de los espectros;  $F_P + F_C - F_M$ , la diferencia de los espectros.

Para cada valor de pH (ver Sección 3.3.2), se monitorizaron los cambios en emisión de fluorescencia que acompañan la titulación fluorimétrica a la longitud de onda donde se presenta el máximo de emisión de la diferencia de los espectros  $F_P + F_C - F_M$ . Las respectivas longitudes de onda de los diversos valores de pH que se realizaron las mediciones de fluorescencia son presentadas en la sexta columna de Tabla 4.4.

Como se puede apreciar en la Figura 4.15, el espectro obtenido de la suma de los espectros de las proteínas en forma libre ( $F_P + F_C$ ) es diferente en magnitud con respecto al espectro de la mezcla ( $F_M$ ). Este apagamiento (también conocido como desactivación o *quenching*) de la señal de emisión de fluorescencia es la evidencia experimental de la formación del complejo. En la Tabla 4.3 se presenta la razón de  $(F_P + F_C)/ (F_M)$ , a los valores de pH en estudio.

**Tabla 4.3** Evidencia experimental de la formación del complejo SCM-quimopapaína–CEW a los diferentes valores de pH en estudio.

pH <sup>I</sup>	$\frac{(F_P + F_C)}{F_M}$	pH <sup>II</sup>	$\frac{(F_P + F_C)}{F_M}$	pH <sup>II</sup>	$\frac{(F_P + F_C)}{F_M}$
<b>3.5</b>	1.19	<b>3.5</b>	1.05	<b>7.6</b>	1.30
<b>4.5</b>	1.60	<b>4.0</b>	1.03	<b>7.8</b>	1.21
<b>5.5</b>	1.67	<b>4.5</b>	1.06	<b>8.0</b>	1.30
<b>6.0</b>	1.88	<b>5.0</b>	1.08	<b>8.5</b>	1.27
<b>6.5</b>	2.21	<b>5.5</b>	1.09	<b>8.8</b>	1.21
		<b>6.0</b>	1.21	<b>9.0</b>	1.15
		<b>6.5</b>	1.22	<b>9.0</b>	1.30
		<b>7.0</b>	1.21	<b>9.2</b>	1.15
		<b>7.2</b>	1.32	<b>9.5</b>	1.12
		<b>7.4</b>	1.17	<b>10.0</b>	1.17
		<b>7.4</b>	1.21		

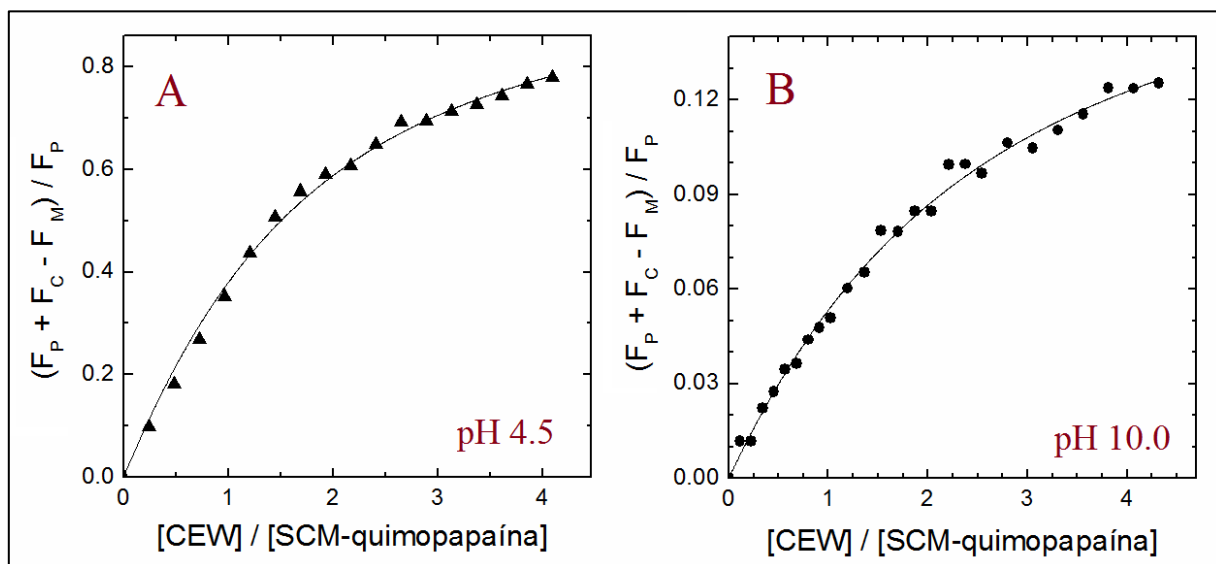
**Donde:** <sup>I</sup> Amortiguador de pH con fuerza iónica de 0.01 M; <sup>II</sup> Amortiguador de pH 50 mM, NaCl 0.1 M;  $F_P$ , fluorescencia de SCM-quimopapaína (0.1  $\mu\text{M}$ );  $F_C$ , fluorescencia de CEW (0.12  $\mu\text{M}$ );  $F_M$ , fluorescencia de la mezcla de SCM-quimopapaína y CEW (0.1  $\mu\text{M}$  y 0.12  $\mu\text{M}$ , respectivamente).

Con propósitos de determinar la constante de disociación mediante el método de titulación fluorimétrica, se requiere una razón de  $(F_P + F_C)/(F_M) > 1.1$  para poder obtener ajustes aceptables, ver coeficiente de determinación ( $r^2$ ) obtenidos en Tabla 4.3.

Cuando se registra  $(F_P + F_C)/(F_M) \leq 1.1$  (ver Tabla 4.3, valores de  $\text{pH}^{\text{II}}$  de 3.5 a 5.5) al evaluar los datos experimentales en la ecuación del modelo matemático (Ecuación 3.2), se registra un coeficiente de determinación ( $r^2$ )  $\leq 0.95$  y se presentan casos donde el ajuste no lineal no converge, resultado de la dispersión de los puntos registrados en la titulación.

Las isotermas presentadas en la Figura 4.16, muestran el comportamiento típico de las titulaciones, a valores de pH ácidos a fuerza iónica de 0.011 M y titulaciones realizadas en el intervalo de pH de 6.0 a 10.0 en presencia de amortiguador 50 mM, NaCl 100 mM. Los resultados de los ajustes a los distintos valores de pH son presentados en la Tabla 4.4.

El perfil de las curvas obtenidas de titulación fluorimétrica de SCM-quimopapaína con CEW (por ejemplo, ver Figura 4.16) son típicas de la titulación de enzima inactiva (bloqueado su sitio catalítico) con cistatina previamente reportadas por Björk & Ylinenjärvi, 1989 [18].



**Figura 4.16** Titulación de SCM-quimopapaína con CEW monitoreado por mediciones de fluorescencia intrínseca de proteínas. Por ejemplo, en presencia de amortiguador de pH 4.5 con fuerza iónica 0.01M (A) y amortiguador de pH 10.0, NaCl 0.1 M (B). Donde:  $F_P$ , fluorescencia de SCM-quimopapaína (0.1  $\mu\text{M}$ );  $F_C$ , fluorescencia de CEW (de 0 a 0.42  $\mu\text{M}$ );  $F_M$ , fluorescencia de la mezcla de SCM-quimopapaína (0.1  $\mu\text{M}$ ) y CEW (de 0 a 0.42  $\mu\text{M}$ ); La línea continua representa el ajuste no lineal de los datos experimentales a al modelo matemático de unión bimolecular reversible (Ecuación 3.2).

**Tabla 4.4** Constante de equilibrio para la unión de SCM-quimopapaína y CEW determinada a diferentes valores de pH a 25 °C.

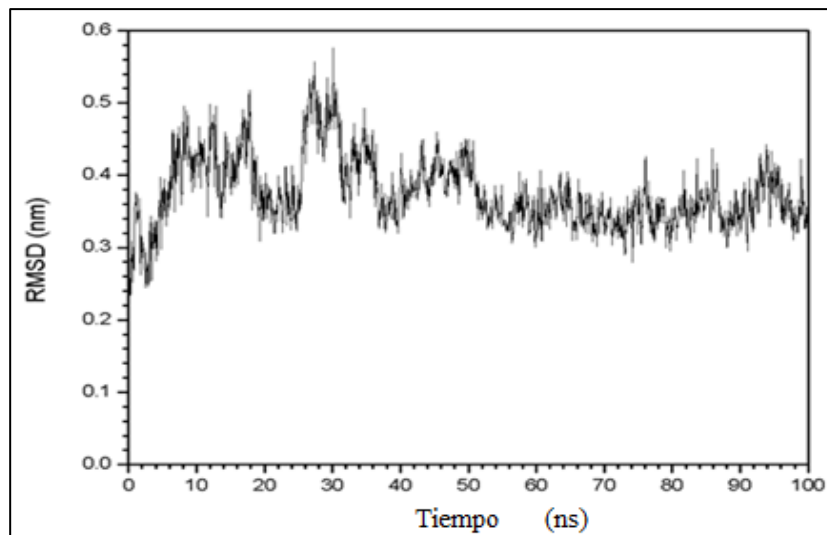
pH	Replicas	$K_b \times 10^{-6} (M^{-1})$	$\alpha$	$r^2$	$\lambda_{\max}$	Amortiguador de pH
Amortiguadores de pH con fuerza iónica de 0.01 M.						
3.5	2	$1.56 \pm 0.03$	$1.47 \pm 0.035$	0.991	338	Formiato
4.5	3	$9.26 \pm 0.29$	$1.05 \pm 0.007$	0.996	338	Acetato
5.5	2	$17.36 \pm 1.36$	$0.82 \pm 0.034$	0.995	340	MES
6.0	3	$45.97 \pm 4.69$	$0.83 \pm 0.018$	0.994	340	MES
6.5	5	$45.08 \pm 2.62$	$0.73 \pm 0.012$	0.960	342	MES
Amortiguador de pH 50 mM , NaCl 0.1 M.						
6.0	3	$4.54 \pm 0.17$	$0.96 \pm 0.047$	0.995	340	MES
6.5	2	$4.87 \pm 0.15$	$0.69 \pm 0.006$	0.992	342	PIPES
7.0	2	$4.13 \pm 0.25$	$0.61 \pm 0.020$	0.992	344	Fosfatos
7.2	2	$4.73 \pm 0.25$	$0.79 \pm 0.016$	0.996	340	Tris
7.4	2	$5.98 \pm 0.49$	$0.46 \pm 0.006$	0.996	342	Fosfatos
7.4	2	$5.44 \pm 0.23$	$0.62 \pm 0.020$	0.993	342	Tris
7.6	2	$5.06 \pm 0.31$	$0.76 \pm 0.010$	0.996	332	Tris
7.8	4	$4.92 \pm 0.23$	$0.51 \pm 0.026$	0.976	342	Tris
8.0	2	$4.20 \pm 0.18$	$0.83 \pm 0.003$	0.994	342	Tris
8.5	3	$4.07 \pm 0.07$	$0.79 \pm 0.020$	0.996	344	Tris
8.8	2	$5.70 \pm 0.16$	$0.34 \pm 0.023$	0.991	342	Tris
9.0	2	$10.24 \pm 0.09$	$0.31 \pm 0.002$	0.995	346	Glicina
9.0	2	$9.87 \pm 0.53$	$0.41 \pm 0.007$	0.990	342	Boratos
9.2	2	$6.85 \pm 0.34$	$0.47 \pm 0.010$	0.987	340	Glicina
9.5	2	$5.92 \pm 0.26$	$0.31 \pm 0.022$	0.989	342	Glicina
10	2	$6.14 \pm 0.73$	$0.18 \pm 0.006$	0.990	346	Glicina

La constante de disociación ( $K_d = K_b^{-1}$ ), constante  $\alpha$  y el coeficiente de determinación ( $r^2$ ) se obtuvieron de la regresión no lineal de las respectivas curvas de titulación, ver Sección 3.3.2. La incertidumbre representa el error estándar de la media (SEM) de las mediciones experimentales.  $\lambda_{\max}$ , indica la longitud de onda donde se registra la máxima diferencia entre la emisión de fluorescencia de las proteínas libres y del complejo.

SEM se calcula por:  $\left(\sigma_M = \frac{\sigma}{\sqrt{N}}\right)$  donde:  $\sigma$  es la desviación estándar de las mediciones,  $N$  indica el número de mediciones.

#### 4.4. Modelo molecular de interacción quimopapaína–CEW.

El modelo del heterodímero quimopapaína–CEW obtenido por reconocimiento molecular (*docking*) fue tomando como estructura de inicio para realizar las simulaciones con MD. El RMSD calculado durante 100 ns de simulación, llegó a un equilibrio después de 52 ns (Figura 4.17).



**Figura 4.17** Evolución de la desviación del valor cuadrático medio (RMSD) estimado para los Ca durante la simulación MD en solvente explícito en un conjunto TNP a 300 K y 1 bar. La estructura de referencia para el cálculo fue la estructura inicial del heterodímero quimopapaína–CEW.

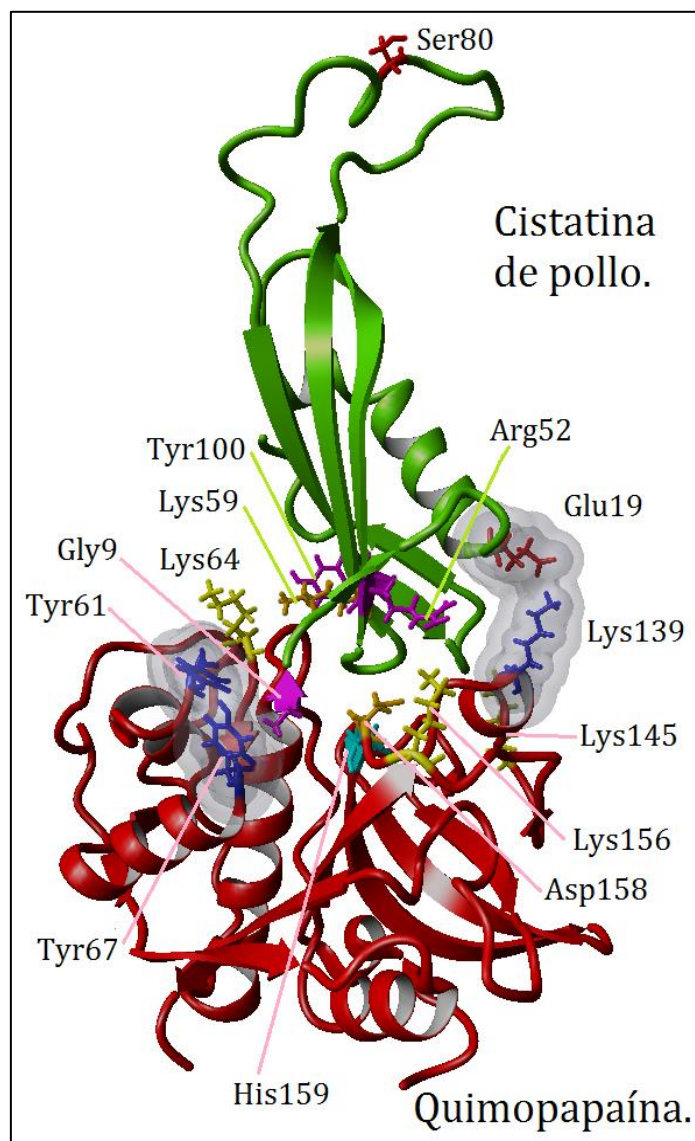
Las interacciones relevantes que exhibe el modelo quimopapaína–CEW en su interfase son presentadas en Tabla 4.5.

**Tabla 4.5.** Interacciones relevantes localizadas en la interfase del modelo quimopapaína–CEW.

Interacciones localizadas en la interfase.	
➤	Seis puentes de hidrógeno.
➤	Un par iónico inter cadena: Lys <sup>139A</sup> [NZ] y Glu <sup>19B</sup> [OE2] (distancia de 2.83 Å).
➤	Dos interacciones catión-Pi: Trp <sup>104B</sup> -Lys <sup>145A</sup> (distancia <sup>catión-Pi</sup> de 8.6 Å, un ángulo de 86.46). Tyr <sup>100B</sup> -Lys <sup>59B</sup> (distancia <sup>catión-Pi</sup> de 5.51 Å, un ángulo de 123.07).
➤	Dos interacciones aromáticas: Trp <sup>181A</sup> - Trp <sup>104B</sup> (distancia <sup>centroide-centroide</sup> de 5.87 Å, un ángulo diedro de 164.39). Trp <sup>185A</sup> - Trp <sup>104B</sup> (distancia <sup>centroide-centroide</sup> de 7.44 Å, un ángulo diedro de 69.09).

Donde: <sup>A</sup>, cadena de quimopapaína; <sup>B</sup>, cadena de CEW.

La estructura final del heterodímero quimopapaína–CEW después de 100 ns de simulación MD se muestra en la Figura 4.18. En esta figura también se ilustran los residuos ionizables que se encontraban dentro de 5.0 Å de cualquier átomo de la cadena opuesta en la interfase.



**Figura 4.18.** Modelo estructural del complejo quimopapaína–CEW. Se muestra los residuos ionizables sobre la interfase del complejo quimopapaína cistatina a un radio de corte <math>< 5 \text{ \AA}</math>, se identifican ocho residuos ionizables sobre quimopapaína ( $D^{158}$ ,  $H^{159}$ ,  $Y^{61}$ ,  $Y^{67}$ ,  $K^{64}$ ,  $K^{139}$ ,  $K^{145}$  y  $K^{156}$ ) y cinco residuos ionizables en cistatina ( $E^{19}$ ,  $Y^{100}$ ,  $K^{59}$ ,  $R^{52}$  y  $G^9_{N\text{-term}}$ ). Figura dibujada con el programa YASARA.



En la Tabla 4.6, se agrupan los residuos localizados en la interfase del heterodímero quimopapaína–CEW de acuerdo en la similitud en el comportamiento de los valores de  $pK_a$  predichos por PROPKA [59] en la proteína en forma separada ( $pK_a^s$ ) y asociada ( $pK_a^c$ ). El cambio en  $pK_a$  al efectuarse la unión de las proteínas puede ser ascendente o descendente. También se muestra el porcentaje de área de superficie accesible al disolvente para cada residuo en las proteínas separadas (% SAS<sup>s</sup>) y en el complejo (% SAS<sup>c</sup>).

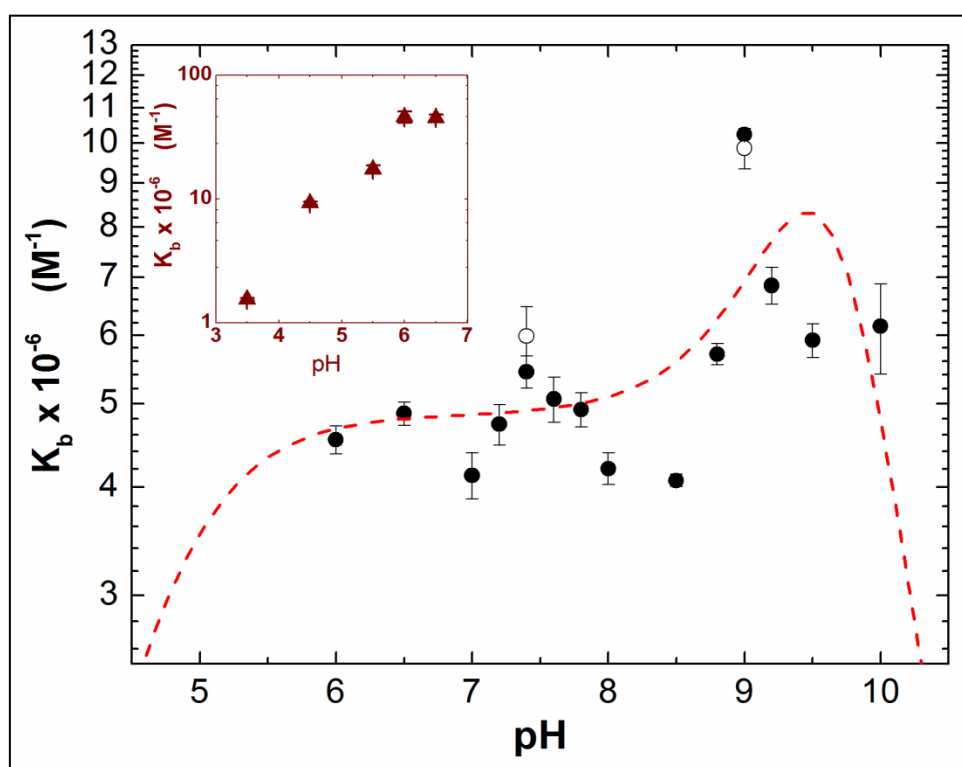
**Tabla 4.6** Propiedades estructurales y fisicoquímicas de residuos ionizables identificados en la interfase del heterodímero quimopapaína–CEW.

Grupo de residuos isoacídicos	Residuo <sup>r</sup>	$pK_a^s/pK_a^c$	% SAS <sup>s</sup> / % SAS <sup>c</sup>	$\Delta G_{b,elec}$ (kJ/mol)	$R^\dagger$
A	Tyr <sup>61A</sup>	11.62/9.58	49.0/40.5	-255.0	0.55
B	Lys <sup>64A</sup>	10.41/10.33	86.5/61.1	-429.5	0.92
	Lys <sup>145A</sup>	10.41/10.23	69.9/51.3	-398.9	0.86
C	Tyr <sup>67A</sup>	9.48/11.61	48.1/39.8	-255.0	0.55
	Lys <sup>139A</sup>	10.46/11.52	76.4/59.1	-339.3	0.73
D	Glu <sup>135A</sup>	3.58/4.3	6.3/6.3	-46.1	0.10
	Asp <sup>158A</sup>	3.04/3.5	60.7/39.6	-46.2	0.10
E	Lys <sup>156A</sup>	11.61/11.28	65.9/59.3	-396.6	0.85
F	His <sup>159A</sup>	6.5/4.2	9.7/2.4	51.8	-0.11
G	Gly <sup>9B</sup>	7.92/6.75	94.6/37.2	-238.4	0.51
H	Glu <sup>19B</sup>	4.6/3.42	73.3/54.5	-104.5	0.22
I	Arg <sup>52B</sup>	13.1/12.93	29.8/22.9	-364.3	0.78
J	Lys <sup>59B</sup>	10.8/8.8	23.6/12.7	-445.5	0.96
K	Tyr <sup>100B</sup>	9.84/11.55	20.1/10.8	-276.4	0.59

Donde: <sup>r</sup>, Residuos ionizables identificados en la interfase, a un radio de corte de 5.0 Å de la cadena opuesta en la interfase del heterodímero quimopapaína–cistatina de pollo; <sup>A</sup>, cadena de quimopapaína; <sup>B</sup>, cadena de cistatina de pollo; <sup>s</sup>, proteína separada; <sup>c</sup>, proteína en el complejo; valores de  $pK_a$  fueron predichos por PROPKA [59]. El porcentaje de área de superficie accesible al solvente (% SAS) para la proteína libre y en complejo se obtuvieron empleando NACCESS [60];  $\Delta G_{b,elec}$ , Energía electrostática de unión para las variantes con cada residuo en el estado protonado alternativo, los cálculos fueron realizados con un modelo de solvente implícito a pH 7.4 y 298 K empleando APBS [61];  $R^\dagger$ , La razón de cambio de  $\Delta G_{b,elec}$ , hace referencia al valor obtenido para el heterodímero nativo a pH 7.4 (-466.0 kJ/mol).

#### 4.5. Dependencia de $K_b$ con respecto a la variación de pH y modelo *MPLM* (Crnogorac *et al.*, 2001).

El gráfico de la constante de unión ( $K_b$ ) con respecto a la variación de pH para la asociación de quimopapaína con cistatina es presentado en la Figura 4.19. La dependencia del pH de la  $K_b$ , exhibe en su perfil un máximo en pH de 9.0 (valor experimental determinado en dos distintos amortiguadores de pH) y tres puntos de inflexión a considerar: El primero en la región ácida de pH, que pudo ser observado solo a baja fuerza iónica (ver inserto de Figura 4.19). El segundo es observado en la región neutra de pH (antes de registrar el máximo). El tercero en la región de pH básico después del máximo.



**Figura 4.19** Efecto del pH sobre la constante de unión ( $K_b$ ) de quimopapaína con CEW. Los experimentos se realizaron en presencia de amortiguador 50 mM, NaCl 0.1 M, 25° C. (ver Tabla 4.4). Las barras en los círculos representan el SEM de las mediciones experimentales. La línea (- - -) representa el mejor ajuste del modelo modificado de unión vinculado a la transferencia de protón (Crnogorac *et al.*, 2001) [34]. El inserto muestra la variación de  $K_b$  para el sistema en valores de pH ácido y una fuerza iónica de 0.011 M.

El análisis de la dependencia de  $K_b$  con respecto a la variación de pH, se realizó usando el modelo *MPLM* [34] desarrollado para la presencia de tres grupos isoacídicos (Ecuación 4.1), correspondiente a los puntos de inflexión observados (ver gráfico presentado en Figura 4.19).

Los valores de  $K_b$  obtenidos en presencia de amortiguador 50 mM, NaCl 0.1 M, 25° C, (presentados en la Tabla 4.4) fueron usados para efectuar el análisis. Se analizaron los perfiles generados por la combinación de tres grupos isoacídicos de la Tabla 4.6, en los respectivos ajustes, los valores de  $m_i$  permanecieron constantes durante el análisis. Los valores de  $pK_a$  del grupo B no presentaron cambios relevantes bajo la unión (este grupo se excluye del análisis).

---

**Ecuación 4.1** Ecuación del modelo *MPLM*, para tres grupos isoacídicos.

---

$$K_b(\text{pH}) = K_{\text{int}}[\text{Num}/\text{Deno}]$$

Donde: *Num* y *Deno* están definidos por:

$$\begin{aligned} \text{Num} = & 1 + 10^{-m_1 pK_{a1}^c} \times 10^{m_1 \text{pH}} + 10^{-m_2 pK_{a2}^c} \times 10^{m_2 \text{pH}} + 10^{-m_3 pK_{a3}^c} \times 10^{m_3 \text{pH}} + \\ & 10^{-(m_1 pK_{a1}^c + m_2 pK_{a2}^c)} \times 10^{(m_1 + m_2) \text{pH}} + 10^{-(m_1 pK_{a1}^c + m_3 pK_{a3}^c)} \times 10^{(m_1 + m_3) \text{pH}} + \\ & 10^{-(m_2 pK_{a2}^c + m_3 pK_{a3}^c)} \times 10^{(m_2 + m_3) \text{pH}} + 10^{-(m_1 pK_{a1}^c + m_2 pK_{a2}^c + m_3 pK_{a3}^c)} \times 10^{(m_1 + m_2 + m_3) \text{pH}} \end{aligned}$$

$$\begin{aligned} \text{Deno} = & 1 + 10^{-m_1 pK_{a1}^s} \times 10^{m_1 \text{pH}} + 10^{-m_2 pK_{a2}^s} \times 10^{m_2 \text{pH}} + 10^{-m_3 pK_{a3}^s} \times 10^{m_3 \text{pH}} + \\ & 10^{-(m_1 pK_{a1}^s + m_2 pK_{a2}^s)} \times 10^{(m_1 + m_2) \text{pH}} + 10^{-(m_1 pK_{a1}^s + m_3 pK_{a3}^s)} \times 10^{(m_1 + m_3) \text{pH}} + \\ & 10^{-(m_2 pK_{a2}^s + m_3 pK_{a3}^s)} \times 10^{(m_2 + m_3) \text{pH}} + 10^{-(m_1 pK_{a1}^s + m_2 pK_{a2}^s + m_3 pK_{a3}^s)} \times 10^{(m_1 + m_2 + m_3) \text{pH}} \end{aligned}$$


---

Esta ecuación es resultado del desarrollo de la ecuación 3.3 (presentado en Apéndice D), permite describir la dependencia de  $K_b$  con respecto a la variación de pH, para tres grupos de residuos isoacídicos. Los parámetros libres a estimar son diez: la constante intrínseca ( $K_{\text{int}}$ ) tres grupos isoacídicos ( $m_1$ ,  $m_2$ ,  $m_3$ , cada uno constituidos de tres posibles residuos), tres valores de  $pK_a$  para la forma asociada ( $pK_{a1}^c$ ,  $pK_{a2}^c$  y  $pK_{a3}^c$ ) y tres valores de  $pK_a$  para la forma libre ( $pK_{a1}^s$ ,  $pK_{a2}^s$  y  $pK_{a3}^s$ ).

---

La única combinación que presenta un perfil semejante al comportamiento de los resultados experimentales (ver, Figura 4.19) fue el obtenido de la combinación de los grupos isoacídicos A, C y H, los cuales poseen 1, 2 y 1 residuos individuales, respectivamente (ver Tabla 4.7). El valor máximo de los datos experimentales registrado a pH de 9.0 no fue reproducido, se registró una aproximación de 86 % respecto al valor medido y un corrimiento hacia la región básica de 0.46 unidades de pH.

Relacionando los parámetros obtenidos del ajuste (presentados en Tabla 4.7) con las propiedades de los residuos que constituyen los grupos isoacídicos de la Tabla 4.6, se aprecian las siguientes conjeturas con respecto a los valores de  $pK_a$ :

- Para el primer grupo isoacídico ( $m_1$ ),  $pK_a^{1s} = 11.90$  y  $pK_a^{1c} = 9.30$  son muy similares a los valores aproximados para Tyr<sup>61A</sup> de 11.62 y 9.58, respectivamente.
- Para el segundo grupo isoacídico ( $m_2$ ), el compuesto de  $pK_a$  incrementa su valor en la forma de la proteína asociada respecto de la proteína en forma libre ( $pK_a^{2s}/ pK_a^{2c} = 9.65/11.9$ ); estos valores son acorde con los predichos para el grupo C formado por Tyr<sup>67A</sup> (9.48/11.61) y Lys<sup>139A</sup> (10.46/11.52).
- También los valores para tercer grupo isoacídico ( $m_3$ ) ( $pK_a^{3s}/ pK_a^{3c} = 4.61/3.39$ ) son concordantes con los predichos para Glu<sup>19B</sup> (4.60/3.42).

**Tabla 4.7** Parámetros obtenidos del mejor ajuste de los datos experimentales de  $K_b$  en función del pH al modelo *MPLM* [34].

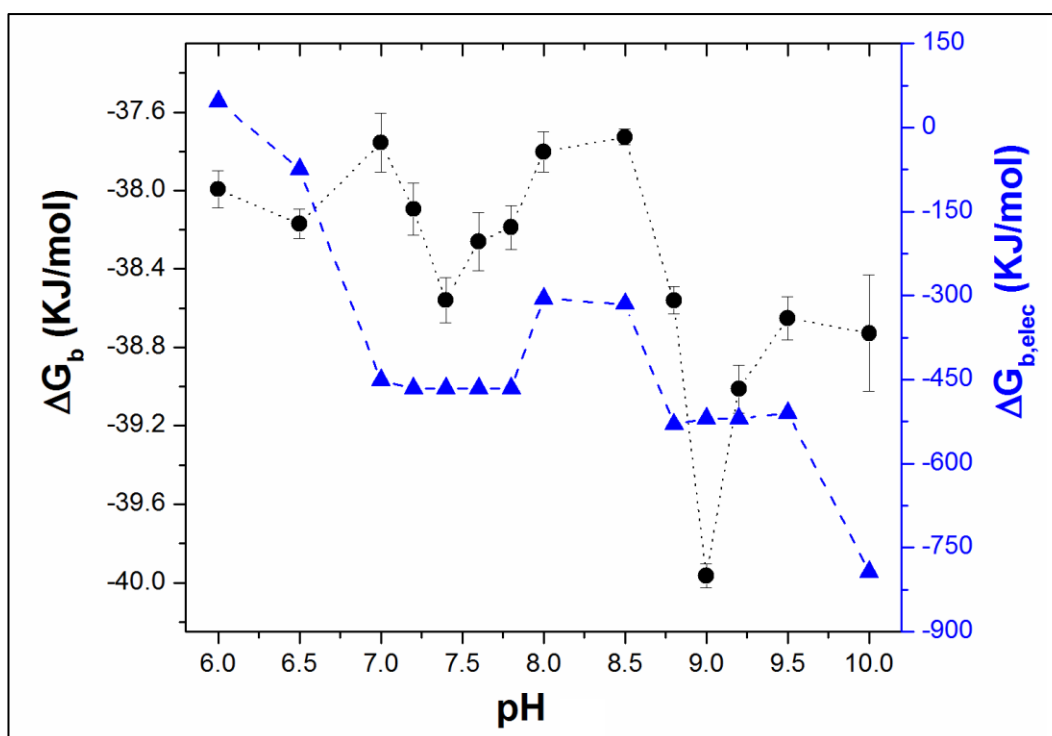
Parámetros					
$m_1$	1	$pK_{a_1}^s$	11.9	$pK_{a_1}^c$	9.3
$m_2$	2	$pK_{a_2}^s$	9.65	$pK_{a_2}^c$	11.9
$m_3$	1	$pK_{a_3}^s$	4.61	$pK_{a_3}^c$	3.39
$K_{int} \times 10^{-6} (M^{-1})$			0.2923		

Donde:  $m_1, m_2, m_3$  denotan la composición del grupo isoacídico, el subíndice indica el grupo isoacídico;  $pK_{a_1}^s, pK_{a_2}^s$  y  $pK_{a_3}^s$  los valores de  $pK_a$  para la forma libre;  $pK_{a_1}^c, pK_{a_2}^c$  y  $pK_{a_3}^c$  son los valores de  $pK_a$  para la forma asociada.

#### 4.6. Comparación de energía libre de unión ( $\Delta G_b$ ) y energía electrostática de unión ( $\Delta G_{b,elec}$ ).

La energía electrostática de unión ( $\Delta G_{b,elec}$ ) para el heterodímero quimopapaína–CEW fue calculada a los diferentes valores de pH en que se determinó la constante de unión ( $K_b$ ), en presencia de los respectivos amortiguadores de pH a una concentración de 50 mM, en presencia de NaCl 0.1 M a 25 °C. También a partir de los valores de  $K_b$  presentados en la Tabla 4.4, se calculó la energía libre de unión ( $\Delta G_b$ ), como:  $\Delta G_b = -RT \ln K_b$ .

Al graficar  $\Delta G_b$  y  $\Delta G_{b,elec}$  versus pH (Figura 4.20), la isoterma de  $\Delta G_{b,elec}$ , describe la tendencia de la isoterma de  $\Delta G_b$ . Estos resultados sugieren que el proceso de unión está regido por interacciones electrostáticas.



**Figura 4.20** Comparación de la energía libre de unión experimental ( $\Delta G_b$ ) -en círculos- y teórico ( $\Delta G_{b,elec}$ ) -en triángulos- del complejo enzima–inhibidor a los diferentes valores de pH examinados. Las barras indican el error estándar de la media (SEM).

La energía de unión electrostática calculada a pH 7.4 y 298.15 K fue  $\Delta G_{b,elec} = -466.0$  kJ/mol. Tomando este valor como referencia, se evaluó la relevancia en la energía de unión de los estados cargados para cada residuo ionizable localizado en la interfase del heterodímero

quimopapaína–CEW. Esta evaluación se realizó cambiando el estado de protonación para cada residuo de forma individual y calculando la energía electrostática de unión. Estos resultados son presentados en la Tabla 4.6 (columnas  $\Delta G_{b,elec}$  y  $R$ ), la última columna hace referencia a la razón ( $R$ ) del valor calculado respecto del valor de referencia,  $\Delta G_{b,elec}$ .

Los residuos que desestabilizan en menor grado al cambiar hipotéticamente su estado de protonación son Lys<sup>59B</sup> y Lys<sup>64A</sup> (registrando un  $R$  de 0.96 y 0.92, respectivamente). Los residuos que se predice ocasionarán mayor desestabilización son Glu<sup>135A</sup> y Asp<sup>158A</sup> e His<sup>159A</sup> (registrando un  $R$  de 0.1, 0.1 y -0.11, respectivamente).

#### 4.7. Comparación estructural de la interacción quimopapaína–CEW con otras estructuras homólogas cisteinproteasas–cistatinas.

##### A. Composición del subsitio $S_2$ de cisteinproteasas comparadas.

El  $S_2$  de quimopapaína está formado por Tyr<sup>67</sup>, Gln<sup>68</sup>, Leu<sup>133</sup>, Leu<sup>157</sup>, Ala<sup>160</sup> y Ser<sup>209</sup> [71]. Es notoria la diferencia en la composición del subsitio  $S_2$  de las enzimas en comparación Tabla 4.8 (Residuos identificados en el alineamiento presentado en la Tabla E.1, Apéndice E).

**Tabla 4.8** Residuos de aminoácidos identificados en el subsitio  $S_2$  de cisteinproteasas.

Enzima	67 <sup>1</sup>	68 <sup>1</sup>	133 <sup>1</sup>	157 <sup>1</sup>	160 <sup>1</sup>	209 <sup>1</sup>
Quimopapaína	Tyr	Gln	Leu	Leu	Ala	Ser
Papaína	Tyr	Pro	Val	Val	Ala	Ser
Catepsina B	Tyr	Pro	Ala	Gly	Ala	Glu
Catepsina H	Leu	Pro	Ala	Val	Ala	Cys
Catepsina L	Leu	Met	Ala	Met	Gly	Ala
Catepsina V	Phe	Met	Ala	Leu	Gly	Ala

<sup>1</sup> Numeración de residuos de aminoácidos en quimopapaína.

La papaína es particularmente buena en hidrolizar enlaces peptídicos que contengan Arg y Lys en el subsitio  $P_1$  (cadena del sustrato), mientras el  $S_2$  de esta enzima es específico a residuos hidrófobicos localizados en subsitio  $P_2$  [8]. Las catepsinas presentan tres subsitios bien definidos  $S_2$  a  $S_1'$  (áreas de contacto con sustrato relativamente pequeña), sin embargo, no son comunes en especificidad de sustratos [8].

### B. Aspectos estructurales de la interacción cisteinproteasas–cistatinas.

Las enzimas comparadas pertenecen a la familia de la papaína (papaína, quimopapaína, catepsinas B, H, L y V), comparten alta homología de secuencia primaria (Tabla E.3, Apéndice E), misma topología global, pero difieren en la cantidad de residuos cargados en su estructura (ver Tabla E.2, Apéndice E). Por su parte, las cistatinas comparadas pertenecen a dos distintas familias (cistatina y estefinas); en consecuencia, comparar la cistatina de pollo con estefinas se observó baja homología de secuencia primaria (estefina A, 15.8 % y estefina B, 21.1%). Respecto a los miembros de la misma familia de estefinas presentan alta homología (Tabla E.3, Apéndice E).

Las cisteinproteasas en comparación son muy similares en composición de estructura secundaria (Tabla E.3, Apéndice E), presentan residuos conservados en su respectivo sitio catalítico, incluyendo tres puentes disulfuro, a excepción de la catepsina B que contiene seis (Tabla E.3, Apéndice E).

Los residuos convenidos en ser esenciales para la actividad catalítica [71] son Cys<sup>25</sup>, Asp<sup>158</sup>, His<sup>159</sup>, Trp<sup>177</sup> y Asn<sup>175</sup>; son todos conservados en las enzimas en comparación (quimopapaína, papaína, catepsina B, H, L y V), a excepción de la sustitución de Asp<sup>158</sup> en catepsina B y H por Gly y Asn, respectivamente (ver alineamiento, Tabla E.1).

### C. Comparación de interfases de complejos homólogos cisteinproteasa–cistatina.

La cantidad de residuos ionizables identificados en la interfase de estructuras homólogas cisteinproteasa–cistatina (Sección 3.4.3) en comparación varía de 11 a 18 residuos (Tabla 4.9), en su mayoría son conservados en las interfaces (Tabla E.1, Apéndice E). En la cadena de la enzima se identificaron de 5 a 11 residuos y por parte del inhibidor, de 4 a 7 residuos. También se identificaron en las distintas interfaces la presencia de 5 a 7 puentes de hidrógeno y de 0 a 2 pares iónicos intercadena. El área de superficie de contacto varía de 1560 a 2640 Å<sup>2</sup>. Cabe mencionar que las proteínas homólogas en comparación presentan diferente constitución de residuos, ver Tabla E.2 (Apéndice E).

**Tabla 4.9** Residuos ionizables<sup>1</sup> y propiedades estructurales de la interfase de heterodímeros homólogos al complejo cisteinproteasa–cistatinas.

Heterodímero	Cisteinproteasa	Cistatina	N <sub>HB</sub>	N <sub>IP</sub>	ASP (Å <sup>2</sup> )
Quimopapaína–CEW <sup>a</sup>	Y <sup>61</sup> , Y <sup>67</sup> , K <sup>64</sup> , E <sup>135</sup> , K <sup>139</sup> , K <sup>145</sup> , K <sup>156</sup> , D <sup>158</sup> , H <sup>159</sup>	G <sup>9</sup> , E <sup>19</sup> , R <sup>52</sup> , K <sup>59</sup> , Y <sup>100</sup>	6	1	1589.3
Papaína–CEW <sup>b</sup>	Y <sup>61</sup> , Y <sup>67</sup> , K <sup>139</sup> , Y <sup>144</sup> , R <sup>145</sup> , K <sup>156</sup> , D <sup>158</sup> , H <sup>159</sup>	G <sup>9</sup> , E <sup>19</sup> , R <sup>52</sup> , K <sup>59</sup> , Y <sup>100</sup>	10	1	1559.9
Papaína–Estefina B <sup>b</sup>	Y <sup>61</sup> , Y <sup>67</sup> , K <sup>139</sup> , R <sup>145</sup> , K <sup>156</sup> , D <sup>158</sup> , H <sup>159</sup>	E <sup>15</sup> , R <sup>68</sup> , H <sup>75</sup> , E <sup>76</sup> , Y <sup>97</sup>	8	1	1965.1
Catepsina B–Estefina A <sup>b</sup>	D <sup>22</sup> , D <sup>69</sup> , Y <sup>75</sup> , E <sup>109</sup> , H <sup>111</sup> , R <sup>116</sup> , D <sup>124</sup> , E <sup>194</sup> , H <sup>199</sup> , D <sup>224</sup> , E <sup>245</sup>	E <sup>8</sup> , K <sup>10</sup> , Y <sup>43</sup> , Y <sup>53</sup> , Y <sup>54</sup> , K <sup>68</sup> , D <sup>79</sup>	13	2	2639.7
Catepsina H–Estefina A <sup>b</sup>	H <sup>64</sup> , R <sup>148</sup> , K <sup>149</sup> , D <sup>162</sup> , K <sup>163</sup> , H <sup>166</sup>	K <sup>10</sup> , E <sup>15</sup> , Y <sup>53</sup> , Y <sup>54</sup> , K <sup>68</sup> , K <sup>71</sup>	9	0	1878.3
Catepsina L–Estefina A <sup>b</sup>	E <sup>141</sup> , D <sup>160</sup> , D <sup>162</sup> , H <sup>163</sup> , E <sup>192</sup>	Y <sup>54</sup> , K <sup>68</sup> , K <sup>71</sup> , E <sup>78</sup>	5	1	1864.4
Catepsina V–Estefina A <sup>b</sup>	K <sup>20</sup> , R <sup>72</sup> , K <sup>148</sup> , K <sup>160</sup> , D <sup>163</sup> , H <sup>164</sup> , E <sup>193</sup>	E <sup>8</sup> , Y <sup>54</sup> , K <sup>68</sup> , K <sup>71</sup>	7	0	1828.1

Donde: <sup>1</sup> Residuos ionizables encontrados dentro de 5.0 Å de cualquier átomo de la cadena opuesta en la interfase. <sup>a</sup> ver Sección 3.4.1; <sup>b</sup> ver Sección 3.4.3; N<sub>HB</sub>, Indica el número de puentes de hidrógeno a través de la interfase; N<sub>IP</sub>, Indica el número de pares iónicos a través de la interfase. ASP, Área de superficie de contacto, calculada como el cambio en el área de superficie accesible al solvente entre el heterodímero y la suma de las proteínas libres usando el programa NACCESS [60].

#### D. Análisis computacional alanine scanning de complejos homólogos cisteinproteasa–cistatina: Residuos hot spots.

Mediante el análisis computacional *alanine scanning*, se exploró la contribución energética en la unión proteína–proteína de residuos individuales localizados en la interfase. La técnica consiste en mutar por Ala cada residuo identificado en la interfase y medir el efecto en la energía libre de unión debida a la hipotética mutación (ver resultados en apéndice F). En la Tabla 4.10 se presenta los residuos que presentan un cambio mayor a 1 Kcal/mol en  $\Delta\Delta G_{\text{bind}}$ , referido como el cambio en la energía de unión de la estructura mutada con respecto a la nativa, estos residuos son nombrados residuos *hot spots* [72,73].



**Tabla 4.10** Residuos *hot spots* identificados en los complejo homólogos cisteinproteasa–cistatina (obtenidos de Tabla F.1, Apéndice F).

Heterodímero	Cisteinproteasa	R <sup>C</sup>	R <sup>NC</sup>	Cistatina	R <sup>C</sup>	R <sup>NC</sup>
Quimopapaína–CEW	Q <sup>19</sup> , Q <sup>142</sup> , H <sup>159</sup> , S <sup>180</sup> , W <sup>181</sup>	9	8	Q <sup>53</sup> , L <sup>54</sup> , V <sup>55</sup> , Y <sup>100</sup> , W <sup>104</sup>	9	2
Papaína–CEW	Q <sup>19</sup> , Q <sup>142</sup> , R <sup>145</sup> , W <sup>177</sup>	12	4	Q <sup>53</sup> , L <sup>54</sup> , V <sup>55</sup> , K <sup>59</sup> , Y <sup>100</sup> , W <sup>104</sup> , N <sup>106</sup>	11	2
Papaína–Estefina B	Q <sup>19</sup> , Y <sup>67</sup> , H <sup>159</sup> , W <sup>177</sup>	17	4	M <sup>2</sup> , Q <sup>46</sup> , V <sup>47</sup> , N <sup>52</sup> , R <sup>68</sup> , L <sup>73</sup> , H <sup>75</sup> , Y <sup>97</sup>	14	3
Catepsina B–Estefina A	Q <sup>23</sup> , N <sup>72</sup> , Y <sup>75</sup> , E <sup>109</sup> , R <sup>116</sup> , E <sup>194</sup> , H <sup>199</sup> , W <sup>221</sup>	20	9	M <sup>1</sup> , I <sup>2</sup> , Q <sup>46</sup> , V <sup>47</sup> , N <sup>52</sup> , F <sup>70</sup> , L <sup>73</sup> , Q <sup>76</sup>	20	4
Catepsina H–Estefina A	Q <sup>67</sup> , L <sup>145</sup> , W <sup>188</sup>	13	8	I <sup>2</sup> , Q <sup>46</sup> , V <sup>47</sup> , N <sup>52</sup> , Y <sup>54</sup> , L <sup>73</sup> , Q <sup>76</sup>	12	2
Catepsina L–Estefina A	Q <sup>19</sup> , L <sup>144</sup> , W <sup>189</sup>	13	5	I <sup>2</sup> , Q <sup>46</sup> , V <sup>47</sup> , N <sup>52</sup> , F <sup>70</sup> , L <sup>73</sup>	11	2
Catepsina V–Estefina A	Q <sup>19</sup> , F <sup>69</sup> , H <sup>164</sup> , W <sup>190</sup>	12	8	I <sup>2</sup> , Q <sup>46</sup> , V <sup>47</sup> , N <sup>52</sup> , F <sup>70</sup> , L <sup>73</sup> , Q <sup>76</sup>	10	2

Donde: R<sup>C</sup>, indica el número de residuos localizados a 4 Å de cualquier átomo perteneciente a la cadena opuesta; R<sup>NC</sup>, los residuos que no tiene contacto pero es oculto bajo la unión.

## 5. DISCUSIÓN.

### 5.1. Purificación de SCM-quimopapaína.

La purificación de quimopapaína se logró en una sola etapa de purificación por cromatografía de intercambio catiónico, la SCM-quimopapaína purificada es muy estable a una concentración de 1 mg/mL, en amortiguador fosfato de sodio 10 mM, pH 7.0. El método de purificación presenta un 10 % de rendimiento, (por ejemplo, si se requiere obtener 1.0 mg de proteína se necesita purificar 10 mg de extracto) lo cual es aceptable en las purificaciones de este tipo de enzimas.

Las propiedades espectrales de la SCM-quimopapaína purificada son características de la proteína, los espectros de fluorescencia y DC (Sección 4.1.2) son acorde a los previamente reportados por otros grupos de investigación [44,66,74]. Respecto a las pruebas de pureza (Sección 4.1.1), la recromatografía muestra la elución de un pico bien definido y simétrico, indicio de que la quimopapaína purificada es homogénea en contenido. El espectro de masas confirma la presencia de una molécula con masa molecular de 23749.3 Da, el valor teórico calculado es de 23752 Da (23638 Da de la secuencia de estructura primaria de quimopapaína más 116 Da correspondientes a la masa de los grupos carboximetil de Cys<sup>25</sup> y Cys<sup>117</sup>).

### 5.2. Purificación de CEW.

La cistatina de pollo se logró purificar con un rendimiento de 8.5 %, con un alto grado de pureza. Por ejemplo, del tratamiento a 1 L de clara (obtenida de dos Kg de huevo, 30 piezas) se obtuvieron aproximadamente 2.1 mg de CEW pura (2.1 mg de CEW pura por lote). Considerando los recursos del laboratorio y principalmente la cromatografía de afinidad en *batch*, el tiempo estimado para procesar un lote fue de un mes.

Esta proteína se intentó obtener en forma recombinante mediante la expresión en *E.coli* por técnicas de Biología Molecular. Sin embargo, no se consiguió dicho objetivo (resultados y avances obtenidos al respecto no presentados).

#### A. Método de purificación de cistatina de pollo a partir de clara de huevo.

La purificación de cistatina de pollo básicamente se realizó en tres etapas (tratamiento de clara de huevo, cromatografía de afinidad y cromatografía por exclusión molecular), suficientes para obtener CEW con alto grado de pureza y presentando actividad inhibitoria a CPs. A continuación presento algunas observaciones a las respectivas etapas de purificación implementadas y respecto a la caracterización de CEW.

a) *Tratamiento de clara de huevo.*

La solubilidad de las proteínas es afectada principalmente por factores como la constitución de aminoácidos, estructura de la proteína (nativa o desnaturalizada), temperatura, contenido de sal (fuerza iónica) y la concentración de proteína [75].

El tratamiento de la clara consiste en modificar la solubilidad de las proteínas contenidas en clara de huevo variando el pH, fuerza iónica y temperatura, con la finalidad de insolubilizar la mayor cantidad de proteína. La remoción de la proteína poco soluble se realizó con precipitación por centrifugación. El tratamiento de la clara se realizó en dos fases:

- En la primera, la solubilidad de las proteínas se afectó con la dilución y adición de NaCl a una concentración final de 0.05 M, y el cambio a pH 3.0; condición reportada por Ferreira *et al.* (2007) [75] donde se registra la más baja solubilidad de proteínas contenidas en la clara. En estas condiciones se modifican principalmente la naturaleza y distribución de carga, provocando cambio en la conformación estructural.
- En la segunda fase, con el tratamiento térmico a 80° C se desestabiliza principalmente a conalbúmina (Temperatura de desnaturalización,  $T_d = 60$  °C), lisozima ( $T_d = 67$  °C), ovoalbúmina ( $T_d = 78$  °C) y S-ovoalbúmina ( $T_d = 82$  °C) [76], favoreciendo la desnaturalización y agregación de estas proteínas.

b) *Cromatografía de afinidad.*

La inmovilización de SCM-papaína sobre sefarosa se realizó con éxito, el método ha sido muy estudiado y no presentó complicaciones. La versatilidad del uso de la cromatografía de afinidad en la purificación de cistatina radica en la estabilidad de ambas proteínas (papaína y cistatina) a pH alcalinos y la alta afinidad que presenta la cistatina por la enzima carboximetilada, para este sistema se ha reportado una  $K_d$  de 0.5 nM ( $K_b = 2 \times 10^9$  M<sup>-1</sup>) a pH 7.4 y una fuerza iónica de 0.15 [18]. La formación del complejo papaína–cistatina es muy estable a los cambios de pH cercanos a la neutralidad, y es fácil desacoplarla a pH de 11.0.

La inmovilización por el método químico (unión covalente al soporte) de SCM-papaína es la etapa del método con mayor importancia, ya que es altamente selectiva. La cromatografía de afinidad es la etapa del proceso en común con los métodos ya reportados (Tabla A.1, Apéndice A). Haciendo una buena cromatografía de afinidad se logra obtener exclusivamente cistatina.

c) *Cromatografía de exclusión molecular.*

La solución de cistatina purificada por el método convencional muestra la presencia de tres bandas, con peso molecular aproximado de 14 KDa, 25 KDa y 47 KDa [51]. Con la columna preparativa únicamente colectamos la proteína con peso molecular aparente de 14 KDa (identificada en gel SDS-PAGE); enseguida, es cargada en una columna de carácter analítico que permite separar las impurezas que pueda contener la muestra de CEW.

El principal problema identificado en la cromatografía de afinidad es el fenómeno de difusión consecuencia de la gran cantidad de proteínas y la baja concentración de cistatina lo cual se podría resolver con un tratamiento adicional que pudiera clarificar aún más la solución (por ejemplo una diafiltración con una membrana de corte de 15 kDa).

En la primera y segunda cromatografía de exclusión molecular sólo se colectó la elución central del pico correspondiente. El rendimiento real de la purificación pudiera mejorarse al recromatografiar las fracciones laterales de elución (colas) con una columna de intercambio aniónico a pH 8.3, en su momento no se contaba con esa columna y buena parte de cistatina no se aprovechó.

También se pudo determinar la composición de la Forma 1 y 2 de cistatina purificada (ver Apéndice B), la composición de las formas depende en gran medida del método de purificación y puede variar de lote a lote [31].

B. *Ensayos de actividad inhibitoria.*

En las preparaciones comerciales de papaína parcialmente purificada están presentes la forma activa e inactiva de la enzima. La CEW forma el complejo enzima-inhibidor con ambas formas. La actividad inhibitoria se probó con un método colorimétrico empleando papaína sin purificar la forma activa de papaína, por lo tanto sólo se puede realizar un análisis cualitativo de la inhibición de la actividad hidrolítica de papaína en presencia de cistatina. Para caracterizar la cinética de hidrólisis de BApNA con papaína, es necesario purificar la forma activa de papaína, lo cual no era el objetivo. Por tal motivo sólo se probó que la cistatina presentara actividad inhibitoria.

### C. *Propiedades espectrales y pruebas de pureza de CEW purificada.*

La CEW purificada presenta propiedades espectrales (espectros de fluorescencia y DC, Sección 4.2.2) previamente reportados por otros grupos de investigación [30].

En relación a las pruebas de pureza (Sección 4.2.1), se observó que la proteína migra en gel SDS-PAGE mostrando una masa molecular aparente próxima a 14 KDa, característica de CEW previamente reportada por otros grupos de investigación [29,40,51]. El espectro de masas de CEW revela la presencia de una molécula con masa molecular de 12161.1 Da y una señal correspondiente a 24314.4 Da. La primera corresponde a CEW cortada en el extremo amino (en Gly<sup>9</sup>) y la segunda a una pequeña señal de fracción dimérica de cistatina (ver Sección 4.2.1).

Abrahamson *et al.* (1987) [19] demostró que al mezclar un exceso molar de CEW o cistatina C con papaína, el complejo enzima-inhibidor es formado inmediatamente. La cistatina libre residual entonces es progresivamente convertida en una forma modificada de cistatina (cortada en el extremo amino en Gly<sup>9</sup>). La inhibición de papaína por cistatina modificada es más de 1000 veces más débil que la cistatina nativa [19].

En la preparación de papaína comercial que utilizaron Abrahamson *et al.* [19], no detectaron contaminación por otra cisteinproteasa (experimentos de cromatografía de intercambio catiónico). El corte en el extremo amino del inhibidor es aparentemente debido a un porcentaje de moléculas cisteinproteasa atípico en la preparación comercial de papaína [19].

Stubbs *et al.* (1990) [22] reportaron que la solución de papaína carboximetilada con la cual trabajaron presenta una actividad residual aproximadamente del 0.5 %. Posiblemente se trate de la misma fracción de papaína comercial atípica, no encontré más información al respecto en la bibliografía revisada.

En el proceso de purificación de CEW, utilicé SCM-papaína en la cromatografía de afinidad. El tiempo de exposición y contacto de SCM-papaína con CEW es de dos horas aproximadamente, una pequeña fracción de papaína comercial atípica pudiera ser la responsable de modificar la CEW (cortada en Gly<sup>9</sup>).

### **5.3. Propiedades espectrales de SCM-quimopapaína y CEW: Espectros de absorbancia en la región UV-Visible, fluorescencia y DC en la región UV-lejano a valores de pH en estudio.**

Los métodos espectroscópicos utilizados (absorbancia, fluorescencia y DC) requieren pocas cantidades de proteína para su análisis y no son destructivos, lo cual permite recuperar la muestra después de realizar el análisis. Se analizaron los cambios en espectros absorbancia, fluorescencia y DC de las proteínas libres (SCM-quimopapaína y CEW) al variar el pH, obteniendo información acerca de su estructura secundaria y terciaria de las proteínas.

Los cambios en los estados de ionización de los grupos polares pueden dar origen a cambios conformacionales en las respectivas estructuras de proteína, los cambios esenciales en composición de estructura secundaria fueron descartados (Sección 4.3.1).

#### *A. Espectros de absorbancia en la región UV-Visible.*

Los espectros de absorbancia obtenidos de SCM-quimopapaína y CEW en la región UV-Visible son típicos de proteína, ambas contienen Trp y Tyr (ver Tabla E.2, Apéndice E) principales responsables de la absorción a 280 nm. En todos los espectros correspondiente a los diferentes valores de pH en estudio, no se presenta el fenómeno batocrómico (desplazamiento hacia el rojo) y está ausente el fenómeno hipsocrómico (desplazamiento hacia el azul).

El efecto hipsocrómico a nivel molecular es un indicio de un aumento en la polaridad del entorno de sus grupos cromóforos, en consecuencia, el efecto batocrómico una disminución de la constante dieléctrica, o bien puede deberse a la presencia del ion tiolato, a valores cercanos de pH 10 ocurre la ionización de la cadena lateral de Tyr:  $\text{TyrOH} \leftrightarrow \text{TyrO}^- + \text{H}^+$  [67,77].

La relevancia de analizar los espectros de absorción en la región UV-Visible radica en el empleo del coeficiente de extinción, el cual se utiliza para calcular la concentración de las proteínas. En el caso hipotético que se presentaran efectos del solvente o bien debidos la ionización de tirosina (ver Sección 4.3.1), sería necesario calcular el coeficiente de extinción bajo esas condiciones (registrando la absorbancia de una disolución de proteína de concentración proteína conocida y usando la ecuación de Lambert-Beer).

### B. Espectros de fluorescencia.

Los cambios en la señal de emisión de fluorescencia dependientes del pH, puede tener su origen por cambios en los estados de ionización de grupos polares en la proximidad de grupos cromóforos Tyr y Trp (a pH ácido los residuos Glu y Asp y a pH alcalino los residuos Tyr, Lys y Arg), provocando una mayor o menor exposición del grupo cromóforo al solvente, desactivación (*quenching*) o propiedades de relajación de las cercanías de los grupos cromóforos. Por ello la importancia de conocer el estado de ionización de los residuos ionizables cercanos a los grupos fluoróforos.

En el caso hipotético que el sistema a estudiar permitiera realizar los experimentos a pH muy básicos, es decir, sin presentar cambios evidentes en estructura secundaria, a valores próximos a pH 10.5–12.0, los cambios en fluorescencia serían causados por la desprotonación de Lys, Tyr y Arg. La gran dificultad de estudiar sistemas que contienen Tyr a pH alcalinos, es considerar la contribución de ion tiosinato; para tal escenario, sería necesario hacer una corrección a las mediciones de fluorescencia.

### C. Espectros de DC en la región UV-lejano

En base a los resultados del análisis de estructura secundaria (Sección 4.3.1) en ambas proteínas, no se detectó cambios relevantes en la composición de estructura secundaria entre los diferentes valores de pH analizados.

Es un hecho que los cambios en estado de ionización de los residuos polares contenidos en la estructura de quimopapaína y CEW (ver Tabla E.2, Apéndice E) den origen a los ligeros cambios registrados en el perfil de los espectros de DC obtenidos a los diversos valores de pH. Sin embargo, en base al análisis de estructura secundaria a distintos valores de pH (Tabla 4.2) no se presentan cambios esenciales en la composición de estructura secundaria.

## 5.4. Determinación de la constante de unión ( $K_b$ ).

La  $K_b$  fue evaluada de acuerdo a una estequiometría de unión bimolecular [30,78], la cual fue demostrada previamente para la interacción quimopapaína–CEW [39]. La concentración para obtener una buena estimación de la constante de afinidad por el método de titulación fluorimétrica es de 1.0 a 0.01  $\mu\text{M}$  [30], con lo cual se pueden determinar  $K_b$  del orden de  $\mu\text{M}$  a nM.

La concentración de quimopapaína utilizada en los experimentos fue de 0.1  $\mu\text{M}$ . Sin embargo, no se pudo determinar la  $K_b$  a valores de  $\text{pH} < 6.0$  en presencia de amortiguador de  $\text{pH}$  50 mM, NaCl 0.1 M (Tabla 4.4), debido a la baja resolución de la señal de fluorescencia. Esto se resuelve aumentando la concentración de la proteína, o bien, bajando la concentración de iones en la solución analizada. Como se contaba con poca cantidad de cistatina, los experimentos en la zona acida de  $\text{pH}$  se obtuvieron bajando la fuerza iónica del amortiguador a 0.01 M (Tabla 4.4).

La diferencia en las magnitud de  $K_b$  determinada a las dos distintas fuerzas iónicas para los valores de  $\text{pH}$  de 6.0 y 6.5 ( $\sim 10$  veces, ver Tabla 4.4) son indicio de que la asociación quimopapaína–cistatina es modulada por interacción electrostática.

El modelo empleado para determinar  $K_d$  sólo requiere del registro de los cambios de emisión de fluorescencia de las proteínas libres y combinadas (Ecuación 3.2) que acompañan la unión molecular. Sin embargo, hay que resaltar la importancia del análisis de los resultados correspondientes a las propiedades espectrales de quimopapaína y CEW (Sección 4.3.1), para descartar interferencias en la determinación de  $K_d$  por el método de titulación fluorimétrica y tener la certeza de que las mediciones de fluorescencia son debidas exclusivamente a la unión molecular.

El modelo estructural del complejo quimopapaína–CEW, exhibe las siguientes interacciones relevantes en la interfase (ver Tabla 4.5), a considerar respecto a las determinaciones de  $K_d$  por medio del método de titulación fluorimétrica: dos interacciones catión-Pi:  $\text{Trp}^{104\text{B}} - \text{Lys}^{145\text{A}}$  y  $\text{Tyr}^{100\text{B}} - \text{Lys}^{59\text{B}}$ . Además, existen dos interacciones aromáticas:  $\text{Trp}^{181\text{A}} - \text{Trp}^{104\text{B}}$  y  $\text{Trp}^{185\text{A}} - \text{Trp}^{104\text{B}}$ . Al respecto de la señal de fluorescencia con la variación de  $\text{pH}$ , las interacciones catión-Pi pudieran relacionarse con el fenómeno de transferencia de carga y las interacciones aromáticas están estrechamente vinculados con un decremento en la señal de fluorescencia por efecto de la exposición y ocultamiento al solvente de Trp.

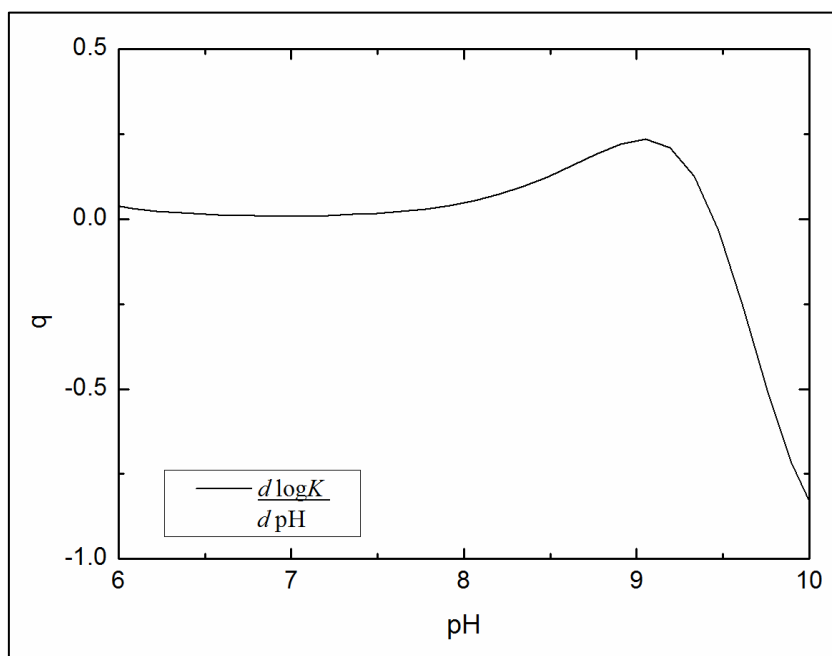
### **5.5. Dependencia de $K_b$ con respecto a la variación de $\text{pH}$ .**

El análisis del efecto del  $\text{pH}$  sobre la constante de unión ( $K_b$ ) de quimopapaína con CEW (Figura 4.19), en combinación de con la identificación de las propiedades estructurales y fisicoquímicas de residuos localizados en la interfase de complejo quimopapaína–CEW (Tabla 4.6), permitieron describir la dependencia de  $K_b$  con la variación del  $\text{pH}$  a través del modelo *MPLM* (Figura 4.19).



Para explicar dependencia de  $K_b$  como función de pH, el modelo de transferencia de protón vinculado a la unión [79,80] implica considerar 16383 parámetros libres para resolver la ecuación resultante (Ecuación D.25, Apéndice D), lo cual es imposible. Con el modelo *MPLM* propuesto por Crnogorac *et al.* (2001), el perfil de los resultados experimentales de  $K_b$  como función de pH se logró describir apropiadamente (Figura 4.19), con la presencia de tres grupos isoacídicos (*composite*  $pK_a$ ), para lo cual sólo se requieren diez parámetros a estimar (ver detalles en Apéndice D).

El modelo *MPLM*, en combinación con valores de  $pK_a$  de residuos ionizables calculados de manera teórica, puede usarse para determinar grupos isoacídicos (*composite*  $pK_a$ ) que actúan concertadamente (o aproximadamente) en la liberación o captación de  $H^+$  [34] bajo algunos cambios del sistema en estudio (Figura 5.1). Esta información no puede ser obtenida por los métodos de NMR y espectroscopia de infrarrojo. Para identificar grupos isoacídicos en una proteína por titulaciones espectroscópicas en función del pH, uno debería realizar una gran cantidad de ensayos [34].



**Figura 5.1** Dependencia del número de iones  $H^+$  ( $q$ ) transferidos hacia la solución amortiguadora de pH con respecto a la variación de pH. La curva representa la solución de la Ecuación 3.4 a los datos obtenidos de la curva del ajuste del modelo *MPLM* presentado en la Figura 4.19. Valores positivos y negativos de  $q$  corresponden a liberación y captación, respectivamente, de iones  $H^+$  bajo la asociación de proteínas [34].

## 5.6. Comparación estructural de la interacción quimopapaína–CEW con otras estructuras homólogas cisteinproteasa–cistatina.

### A. Composición del subsitio $S_2$ de cisteinproteasas comparadas.

Respecto a la reactividad de CPs por sus sustratos, Turk *et al.* (1997) aseveran que la especificidad de estas enzimas parece ser gobernada principalmente por la estructura de los subsitios  $S_2$  a  $S_3'$  [81]. La constitución de estos subsitios restringen la accesibilidad del sustrato sobre la hendidura de unión de la enzima, esta topografía complementaria del sustrato al sitio activo de la enzima parece ser la clave de la diferencia en el mecanismo catalítico de CPs. La actividad catalítica de CPs puede clasificarse en endopeptidasa (por ejemplo: catepsina L, S, V y F) o exopeptidasa (por ejemplo: catepsina B, C, H y X) [8,81]. El grado de flexibilidad de conformación que pueda tomar la región cercana al sitio activo de la enzima también es importante para la fijación del sustrato sobre la enzima (teoría del acoplamiento inducido de la acción enzimática).

Maes *et al.* (1996) examinaron y compararon la hendidura del sitio activo de quimopapaína, concluyen que no difiere significativamente en comparación con papaína, caricaína y glicilendopeptidasa (enzimas contenidas en látex de *Carica papaya*) [12]. Sin embargo, la sustitución de los residuos Val<sup>133</sup> y Val<sup>157</sup> de papaína (localizados en el subsitio de  $S_2$ ) por los residuos de cadena lateral más grande Leu<sup>133</sup> y Leu<sup>157</sup> en quimopapaína explican la diferencia catalítica entre estas enzimas [71]. Esta especificidad relativa se explica por el remplazo de estos residuos en el subsitio de  $S_2$  de quimopapaína, que estructuralmente restringe la accesibilidad de sustratos que contienen Phe en el subsitio  $P_2$  [82].

### B. Aspectos estructurales de la inhibición de cisteinproteasas por cistatinas.

El sitio de unión de la cisteinproteasa con la cistatina [21] y estefinas [22] está compuesto por dos asas (hairpin loops) y un segmento amino terminal (por ejemplo, ver Figura 1.5). La primera asa contiene la secuencia conservada QVVAG y en la segunda asa se encuentra Pro-Trp se encuentran enlazados en la zona  $S_1'$ - $S_2'$ , en tanto el amino terminal del inhibidor el cual contiene el residuo Gly<sup>9</sup> conservado en todas las cistatinas interacciona en los sitios  $S_1$ - $S_3$  de la enzima. En conjunto forman una cuña hidrofóbica que es altamente complementaria al sitio activo arquetipo de la familia de CPs permitiendo una inhibición extremadamente fuerte, rígida y muy rápida [8,21].

El mecanismo de inhibición no involucra el residuo catalítico Cys<sup>25</sup> (inhibición exosítio), el cual se encuentra lejano al segmento amino terminal para interactuar (una distancia de 7.49 Å entre Cys<sup>25</sup> [SH] de quimopapaína y GLY<sup>9</sup> [N] de CEW).

Turk *et al.* (1997) [81] argumentaron que las diferencias en la constante de unión de cistatinas y varias enzimas, surge principalmente de las diferencias en la estructura de la hendidura del sitio activo. Por ejemplo, la inhibición de endopeptidasas tales como papaína, catepsina L, S y cruzipapaína por cistatinas de origen animal es extremadamente fuerte y rápida [81], mientras la inhibición de exopeptidasas tales como catepsina B y catepsina H es considerablemente más débil [81] (por ejemplo, ver magnitudes en Tabla 1.5). Al respecto, se conoce que la hendidura del sitio activo de las endopeptidasas es libre de acceder y acoplarse el inhibidor, mientras que en el caso de las exopeptidasas la hendidura del sitio activo contiene residuos adicionales de la enzima que restringe libertad de acoplarse el inhibidor.

A pesar de similitud del mecanismo de reacción de cistatinas y de las estructuras de interacción de cistatinas con cisteinproteasas tipo papaína, hay diferencias importantes en la inhibición individual de enzimas. Debido a que las interacciones entre inhibidores y enzimas son en su mayoría no específicas [22].

*C. Análisis computacional alanine scanning de complejos homólogos cisteinproteasa–cistatina: Residuos hot spots.*

Como se puede apreciar en la Tabla F.1 (Apéndice F), los residuos identificados en la interfase de los complejos analizados que hace una importante contribución a la energía libre de unión, son residuos conservados en cisteinproteasa y cistatinas respectivamente.

*D. Características de la interacción de estructuras homólogas cisteinproteasa–cistatina.*

Son pocos los estudios reportados acerca de la interacción enzima–inhibidor con respecto a la variación del pH. En el caso de la interacción cisteinproteasa–cistatina, solo está reportado un estudio de papaína con CEW [39,42,43], a pesar de los avances realizados en este campo los resultados no son concluyentes para establecer un mecanismo general de interacción de cisteinproteasa–inhibidor cistatina. La magnitud de la  $K_d$  de estos sistemas varía del orden de  $10^{-6}$  M ( $\mu$ M) a  $10^{-15}$  M (fM). Las características generales de la interacción de estructuras homólogas cisteinproteasa–cistatina, identificadas en el presente trabajo son presentadas en la Tabla 5.1.

**Tabla 5.1** Características generales de la interacción de estructuras homólogas cisteinproteasa–cistatina.

Característica	Observación
1. La cantidad de residuos ionizables varía de 11 a 18 (en enzima de 5 a 11 y en inhibidor de 4 a 7).	En su mayoría son residuos conservados. La estructura y estabilidad del complejo dependerá del estado de ionización de los residuos individuales y de sus posibles contactos con otros residuos.
2. La presencia puentes de hidrógeno es de 5 a 7 puentes de hidrógeno.	No están considerados los enlaces con moléculas de aguas conservadas, localizadas en la interfase del respectivo complejo.
3. El área de superficie de contacto varia de 1560 a 2640 Å <sup>2</sup> .	El área de contacto no es determinante en la magnitud de $K_b$ .
4. Pares iónicos intercadena de 0 a 2.	Se ha visto que la presencia de este tipo de interacción, estabiliza la formación del complejo y aumenta la magnitud de la contante de $K_i$ [83].
5. El número de residuos <i>hot spots</i> varía de 10 a 18 (en enzima de 3 a 8 y en inhibidor de 5 a 8).	Son residuos conservados.

Donde: RC, indica el número de residuos localizados a 4 Å de cualquier átomo perteneciente a la cadena opuesta; RNC, los residuos que no tiene contacto pero es oculto bajo la unión.

## 6. CONCLUSIONES.

- 1) La constante de la unión ( $K_b$ ) de quimopapaína -una cisteinproteasa representativo tipo papaína- con el inhibidor cistatina de pollo (CEW) muestra dependencia con respecto a la variación del pH y fuerza iónica ( $I$ ).
- 2) La unión de SCM-quimopapaína con cistatina de pollo (en su forma cortada en el extremo amino terminal en Gly<sup>9</sup>) presenta una moderada afinidad.
- 3) La SCM-quimopapaína y cistatina de pollo (ambas en su forma libre) en el intervalo de pH de [3.5–10.0] a 25° C, no presentan un cambio significativo en la composición de estructura secundaria de proteína y está ausente efectos de la presencia de la ionización de Tyr en las respectivas señales espectrales.
- 4) La asociación quimopapaína–CEW es modulada por interacción electrostática. Estas interacciones son moduladas por el pH del entorno, que dicta el estado de protonación de los residuos ionizables localizados en la interfase del heterodímero de acuerdo con su valor de  $pK_a$  en la forma asociada o separada de las proteínas, por lo tanto, también determina la energía de unión electrostática.
- 5) Proponemos que los residuos que participan en el par iónico entre cadenas, Lys<sup>139</sup> de quimopapaína y Glu<sup>19</sup> de CEW, así como Tyr<sup>61</sup> y Tyr<sup>67</sup> de quimopapaína, son los principales residuos responsables de la dependencia de la afinidad del heterodímero quimopapaína–CEW observada experimentalmente con respecto a la variación del pH.
- 6) La variación de  $K_b$  (calculada como  $K_b=K_d^{-1}$ ) con el pH se logró describir apropiadamente con el modelo modificado de transferencia de protón (*MPLM*), propuesto por Crnogorac *et al.* (2001), con la presencia de tres grupos ionizables isoacídicos.
- 7) La dependencia de la constante de unión con respecto a la variación del pH, así como la contribución electrostática a la energía de unión están fuertemente vinculada a la presencia y distribución de grupos ionizables en la interfase de complejos homólogos cisteinproteasa–cistatina.

## 7. PERSPECTIVAS.

- A) Determinar la  $K_d$  del sistema quimopapaína-cistatina de pollo a cinco temperaturas (20, 25, 30, 35 y 40 °C) para poder determinar el cambio en la entalpia de unión ( $\Delta H_b$ ) y el cambio en la capacidad calorífica de unión ( $\Delta C_{p_b}$ ), mediante el uso de la ecuación integrada de van't Hoff. Los experimentos se realizarían con cistatina de pollo obtenida a partir de clara de huevo, y empleando el método de titulación fluorimétrica.
- B) Expresar y purificar cistina de pollo recombinante para poder realizar experimentos de Calorimetría de Titulación Multitermica (MTC) que permitiera hacer una caracterización termodinámica del sistema quimopapaína–cistatina de pollo.

## 8. REFERENCIAS.

1. Berg, J. M.; Tymoczko, J. L.; Stryer, L. Bioquímica, 5ª Edición, *Editorial Reverte*, **2004**, pág.: 41-73, 84, 85, 188-192, 237.
2. Feijoo-Siota, L.; Villa T. G. Native and biotechnologically engineered plant proteases with industrial applications. *Food Bioprocess Technology*, **2011**, 4, 1066-1088.
3. Baines, B. S.; Brocklehurst, K.; Isolation and characterization of the four major cysteine proteinase components of the latex of *Carica papaya* L. Reactive characteristics towards 2'2-dipyridyl disulfide of the thiol groups of papain, chymopapains A and B, and papaya peptidase A. *Journal of Protein Chemistry*, **1982**, 1(2), 119-139.
4. Oliveira, A. S.; Filho, J. X. ; Sales P. Cysteine proteinases and cistatins. *Brazilian Archives of Biology and Technology an International Journal*, **2003**, 46, 1, 91 - 104.
5. Dubey, V. K; Pande, M. S.; Jagannadham, M. V. Papain-like proteases: Applications of their inhibitors. *African Journal of Biotechnology*, **2007**, 6(9), 1077-1086.
6. Grzonka, Z.; Jankowska, E.; Kasprzykowski, F.; Kasprzykowska, R.; Łankiewicz, L.; Wiczak, W.; Wieczerek, E.; Ciarkowski, J.; Drabik, P.; Janowski, R.; Kozak, M.; Jaskólski, M.; Grubb, A. Structural studies of cysteine proteases and their inhibitors. *Acta Biochimica Polonica*, **2001**, 48, 1-20.
7. Rawlings, N. D.; Morton, F. R.; Barrett, A. J. In J. Polaina and A.P. MacCabe (Eds.): *Industrial Enzymes*, Springer. **2007**; An introduction to peptidases and the MEROPS database, pp. 161–179.
8. Otto, T. T.; Schirmeister, T. Cysteine proteases and their inhibitors. *Chemical Reviews*, **1997**, 97,133-171.
9. Rzychon, M.; Chmiel, D.; Stec-Niemczyk, J. Modes of inhibition of cysteine proteases. *Acta Biochimica Polonica*, **2004**, 51(4), 861-873.
10. Sumner, I. G.; Harris, G. W.; Taylor M. A. J.; pickersgill, R. W.; Owen, A. J.; Goodenough, P. W. Factors effecting the thermostability of cysteine proteinases from *Carica papaya*. *European Journal of Biochemistry*, **1993**, 214, 129-134.
11. Buttle, D. J.; Barret A. L. Chymopapain: Cromatographic purification and immunological characterization. *Biochemical Journal*, **1984**, 223, 81-88.
12. Maes, D.; Bouckaert, J.; Poormans, F.; Wyns, L.; Looze, Y. Structure of Chymopapain at 1.7 a resolution. *Biochemistry*. **1996**, 35, 16292-16298.
13. Gomes, M. T. R.; Teixeira, R. D.; Lopes, M. T. P.; Nagen R. A. P.; Salas, C. E. X-ray cristal structure of CMS1MS2: a high proteolytic activity cysteine proteinase from *Carica candamarcensis*. *Amino Acids*. **2012**, 43, 2381-2391.

14. Krieger, E.; Koraimann, G.; Vriend, G. Increasing the Precision of Comparative Models with YASARA NOVA—a Self-Parameterizing Force Field. *PROTEINS: Structure, Function, and Genetics*, **2002**, 47, 393-402.
15. Jacquet, A.; Kleinschmidt, T.; Schnek, A. G.; Looze, Y.; Braunitzer, G.; The tiol proteinases from the latex of *Carica papaya L.* III. The primary structure of chymopapain. *Biological chemistry Hoopse-Seyler*, **1989**, 370, 425-434.
16. Trziszka, T.; Saleh, Y.; Kopeć, W.; Siewiński, M.; Węsierska, E. Effect of hen's age on the level of cystatin in the chicken egg white. *International Journal of Poultry Science*, **2004**, 3(7), 471 – 477.
17. Rawlings, N. D.; Tolle, D. P.; Barrett, A. J. Evolutionary families of peptidase inhibitors. *Biochemical Journal*, **2004**, 378, 705–716.
18. Björk, I. and Ylinenjärvi, K. Interaction of chicken cystatin with inactivated papains. *Biochemical Journal*, **1989**, 260, 61 - 68.
19. Abrahamson, M.; Ritonjasl, A.; Brown, M. A.; Grubb, A.; Machleidts, W.; Barrett, J.A. Identification of the probable inhibitory reactive sites of the cysteine proteinase inhibitors human cystatin C and chicken cystatin. *The Journal of Biological Chemistry*, **1987**, 262, 9688-9694.
20. Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Research*, **2000**, 28(1), 235-42.
21. Bode, W.; Engh, R.; Musul, D.; Thiele, U.; Huber, R.; Karshikov, A.; Brzin, J.; Kos, J.; Turk, V. The 2.0 Å X-ray crystal structure of chicken egg white cystatin and its possible mode of interaction with cysteine proteinases. *The EMBO Journal*, **1988**, 7(8), 2593 – 2599.
22. Stubbs, M. T.; Laber, B.; Bode, W.; Huber, R.; Jerala, R.; Lenarcic, B.; Turk, V. The refined 2.4 Å X-ray crystal structure of recombinant human stefin B in complex with the cysteine proteinase papain: a novel type of proteinase inhibitor. *The EMBO Journal*, **1990**, 9(6), 1939-1947.
23. Kovacs-Nolan, J.; Phillips, M.; Mine, Y. Advances in the value of eggs and egg components for human health. *Journal of Agricultural and Food Chemistry*. **2005**, 53, 8421 - 8431.
24. Carraro, A. A. C. Albumen protein and functional properties of gelation and foaming. *Sci. Agric. (Piracicaba, Braz)*. **2006**, 63 No.3 Piracicaba May/june.



25. Schwabe, C.; Anastasi, A.; Crow, H.; Mc Donald, J. K.; Barrett, A. J. Cystatin: Aminoacid sequence and posible secondary structure. *Biochemical Journal*, **1984**, 217, 813-817.
26. Zerovnik, E.; Cimermen, N.; Kos, J. Turk V.; Lohner, K. Thermal Denaturation of human cystatin C and two of its variants; comparison to chicken cystatin. *Biological chemistry*, **1997**, 378, 1199-1203.
27. Barret, A. J. Cystatin, the egg white inhibitor of cisteine proteinases. *Methods in Enzimology*, **1981**, 80, 771 - 778.
28. Réhault, S. INRA, HR 83 Recherches Avicoles, F-37380 Nouzilly, France. Antiproteases. Bioactive Egg Compounds (ed. By R.Huopalahti *et al.*) © Springer-Verlag Berlin Heidelberg **2007**.
29. Anastasi, A.; Brown, M. A.; Kembhavi, A. A.; Nicklin, M. J.; Sayers, C.A.; Sunter, D. C.; Barrett, A. J. Cystatin, a protein inhibitor of cisteine proteinases. *Biochemical Journal*, **1983**, 211, 129 – 138.
30. Lindahl, P.; Alriksson, E.; Jörnvall, H.; Björk, I. Interaction of the cysteine proteinase inhibitor chicken cystatin with papain. *Biochemistry*, **1988**, 27, 5074 – 5082.
31. Laber, B., Krieglstein, K.; Henschen, A.; Kos, J.; Turk, V.; Huber, R.; Bode, W. The cysteine proteinase inhibitor chicken cystatin is a phosphoprotein. *FEBS*, **1989**, 248(1-2), 162 - 168.
32. Nicklin, M. J. H.; Barret, A. J. Inhibition of cysteine proteinases and dipeptidyl peptidase I by egg-white cystatin. *Biochemical Journal*, **1984**, 223, 245-253.
33. Dong, F.; Olsen, B.; Baker, N.A. Computational methods for biomolecular electrostatics. *Methods in Cell Biology*, **2008**, 84, 843-70.
34. Crnogorac, M. M.; Ullmann, G. M.; Kostić, N. M. Effects of pH on Protein Association: Modification of the Proton-Linkage Model and Experimental Verification of the Modified Model in the Case of Cytochrome c And Plastocyanin. *Journal of the American Chemical Society*, 2001, 123, 10790-10798.
35. Jensen, J. H. Calculating pH and salt dependence of protein-protein binding. *Current Pharmaceutical Biotechnology*, **2008**, 9(2), 96-102.
36. Kukić, P.; Nielsen, J.E. Electrostatics in proteins and protein-ligand complexes. *Future Medicinal Chemistry*, **2010**, 2(4), 647-66.
37. Kirschke, H.; Barret, A. J.; Rawlings N.D. Lysosomal Cysteine Proteases. *Published in The United States by Oxford University Press, Inc. New York*, **1998**; 54-57.

38. Bode, B.; Huber, R. Structural basis of the endoproteinase-protein inhibitor interaction. *Biochimica et Biophysica Acta*, **2000**, 1477, 241-252.
39. Björk, I.; Ylinenjärvi, K. Interaction between chicken cystatin and the cysteine proteinases actinidin, chymopapain A, and ficin. *Biochemistry*, **1990**, 29(7), 1770-1776.
40. Barrett, A. J. The cystatins: a new class of peptidase inhibitors. *Trends in Biochemical Sciences*, **1987**, 12, 193-196.
41. Klemenčič, I.; Carmona, A. K.; Cezari M. H. S.; Juliano, M. A.; Juliano L.; Gunčarg, G.; Turk, D.; Krizčaj, I.; Turk, V.; Turk B. Biochemical characterization of human cathepsin X revealed that the enzyme is an exopeptidase, acting as carboxymonopeptidase acting as carboxymonopeptidase or carboxydiptidase. *European Journal of Biochemistry*, **2000**, 267, 5404-5412.
42. Lindahl, P.; Abrahamson, M.; Björk, I. Interaction of recombinant human cystatin C with the cysteine. *Biochemical Journal*, **1992**, 281, 49-55.
43. Björk, I.; Alriksson, E.; Ylinenjärvi, K. Kinetics of binding of chicken cystatin to papain. *Biochemistry*, **1989**, 28(4), 1568-1573.
44. Solis-Mendiola, S.; Arroyo-Reyna, A.; Hernández-Arana, A. Circular dichroism of cysteine proteinases from papaya latex. Evidence of differences in the folding of their polipeptide chains. *Biochimica et Biophysica Acta*, **1992**, 1118, 288-292.
45. Azarkan, M.; Dibiani, R.; Goormaghtigh, E.; Raussens, V.; Baeyens-Volant, D. The papaya kunitz-type trypsin inhibitor is a highly stable  $\beta$ -sheet glycoprotein. *Biochimica et biophysica Acta*, **2006**, 1764: 1063 – 1072.
46. Levine, R. L.; Federici, M. M. *Quantitation of aromatic residues in proteins: Model compounds for second-derivative spectroscopy*. *Biochemistry*, **1982**, 21, 2600-2606.
47. Kelly, S. M.; Jess, T. J.; Price, N. C. How to study proteins by circular dichroism. *Biochimica et Biophysica Acta*, **2005**, 1751, 119 – 139.
48. Turk, V.; Brzin, J.;Longer, M.; Ritonja, A.; Eropkin, M.; Borchart, U.;Machleidt,W. Protein inhibitors of cysteine proteinases. III. Amino-acid sequence of cystatin from chicken egg white. *Hoppe Seylers Zeitschrift fur physiologische Chemie*, **1983**, 364(11), 1487-1496.
49. Turk, V.; Brzin, J. Process for the isolation of chicken egg cystatin, antiviral agents containing it and its use as viral protease inhibitor US 4902509 A. *United States Patent*, PN: 4902509, DP: Feb.20 **1990**.

50. Trziszka, T. ;Polanowski, A. ;Saleh, Y.;Kopec, W. World's Poultry Science Association. Submission ID 10776, Category E.5. Egg white cystatin isolation and its medical application. In *WORLDS POULTRY SCIENCE JOURNAL -WYTON, European poultry conference 12th, European poultry conference, 2006*, 62, 170.
51. Wang, J.; Wu, J. An improved method to extract and purify cystatin from hen egg white. *J. Chromatogr. B Analyt Technol Biomed. Life Sci.* **2014**, 963, 10-5.
52. Zerovnik, E.; Cimernen, N.; Kos, J.; Turk V.; Lohner, K. Thermal Denaturation of human cystatin C and two of its variants; Comparison to chicken cystatin. *Biological chemistry*, **1997**, 378, 1199-1203.
53. Saxena, I.; Tayyab, S. Protein proteinase inhibitors from avian egg whites. *CMLS: Cellular and Molecular Life Sciences*, **1997**, 53, 13 - 23.
54. Jimenez-García, B.; Pons, C.; Fernández-Recio, J. pyDockWEB: a web server for rigid-body protein-protein docking using electrostatics and desolvation scoring. *Bioinformatics*, **2013**, 29, 1698-1699.
55. Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *Journal of Chemical Theory and Computation*, **2008**. 4, 435–447.
56. Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *Journal of the American Chemical Society*, **1996**, 118, 11225–11236.
57. Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije J. G. E. M. LINCS: a linear constraint solver for molecular simulations. *Journal of Computational Chemistry*, **1997**, 18, 1463–1472.
58. Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An  $N \cdot \log(N)$  method for Ewald sums in large systems. *Journal of Physical Chemistry*, **1993**, 98, 10089–10092.
59. Li, H.; Robertson, A. D.; Jensen, J. H. Very fast empirical prediction and interpretation of protein pKa values. *Proteins*, **2005**, 61, 704-721.
60. Hubbard, S. J.; Thornton, J. M. 'NACCESS', Computer Program, Department of Biochemistry and Molecular Biology, University College London. **1993**.  
<http://www.bioinf.manchester.ac.uk/naccess/>
61. Baker, N. A.; Sept, D.; Joseph, S.; Holst, M. J.; McCammon, J. A. Electrostatics of nanosystems: Application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, 98, 10037-10041.

62. Lira-De León, K. I.; García-Gutiérrez, P.; Serratos, I. N.; Palomera-Cárdenas, M.; Figueroa-Corona, M. P.; Campos-Peña, V.; Meraz-Ríos, M. A. Molecular mechanism of Tau aggregation induced by anionic and cationic dyes. *Journal of Alzheimer's Disease*, **2013**, 35, 319–334.
63. Sitkoff, D.; Sharp, K. A.; Honig, B. Accurate calculation of hydration free energies using macroscopic solvent models. *Journal of Physical Chemistry*, **1994**, 98, 1978–1988.
64. Dolinsky, T.J.; Nielsen, J. E.; McCammon, J. A.; Baker, N. A. PDB2PQR: an automated pipeline for the setup, execution, and analysis of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Research*, **2004**, 32, W665-W667.
65. Krieger, E.; Joo K.; Lee, J.; Lee, J.; Raman, S.; Thompson, J.; Tyka, M.; Baker, D.; Karplus, K. Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: Four approaches that performed well in CASP8. *Proteins*, **2009**, 77, 114-22.
66. Solis-Mendiola, S.; Zubillaga-Luna, R.; Rojo-Domínguez, A.; Hernández-Arana. Structural similarity of chymopapain forms as indicated by circular dichroism. *Biochemical Journal*, **1989**, 257,183-186.
67. Kunimitsu, Donald Kunio. Thesis Ph. D.: The purification and some physical-chemical studies of cristaline chymopapain B. **1964**, University of Hawaii, Chemistry, biological.
68. Deléage, G.; Geourjon, C. An interactive graphic program for calculating the secondary structure content of proteins from circular dichroism spectrum. *Computer Applications in the Biosciences*, **1993**, 9(2), 197-199.
69. Schmid, F. X. Spectral methods of characterizing protein conformation and conformational changes, In Protein Structure: a practical approach (2nd edition) (edited by T. Creighton). *Oxford University Press, Oxford*, **1996**, 251-285
70. Campbell, I. D.; Dwek, R. A. Fluorescence Chapter 5, In Biological spectroscopy, *Benjamin/Cumminngs Publishing*, Menlo Park, C.A., **1984**, 91-108.
71. Watson, D. C.; Yaguchi, M.; Lynn, K. R. The amino acid sequence of chymopapain from Carica papaya. *Biochemical Journal*, **1990**, 266(1), 75-81.
72. Kortemme, T.; Baker, D. A simple physical model for binding energy hot spots in protein-protein complexes. *Proceedings of the National Academy of Sciences U.S.A*, **2002**, 99(22), 14116-14121.
73. Kortemme, T.; Kim, D. E.; Baker, D. Computational alanine scanning of protein–protein interfaces. *In Science Signaling, PROTOCOLS*, **2004**, 1-8.
74. Barely, A. O.; Glazer, A. N. Spectroscopic studies on papain and some inactive derivatives. *The Journal of Biological Chemistry*, **1969**, 244(2), 268-273.

75. Ferreira, M. F.; Coimbra, J. S. R.; Garcia, R. E. E.; Minim, L. A.; Oliveira, F. C.; Sousa, R. C. S. Solubility and density of egg white proteins: Effect of pH and saline concentration. *LWT - Food Science and Technology*, **2007**, 40(7), 1304 - 1307.
76. Ferreira, M.; Hofet, C.; Raemy, A. A calorimetric study of egg white proteins. *Journal of Thermal Analysis*, **1997**, 48,683-690.
77. Garcia-Segura, J. M.; Gavilanes, J. G.; Martínez del pozo, A.; Montero, F.; Oñaderra, M.; Vivanco, F. Técnicas instrumentales de análisis en bioquímica. **2002**, Editorial Síntesis, 89-95.
78. Reyes-Espinosa, F. Determinación de la constante de disociación para el sistema S-(carboximetil)-caricaína-cistatina. Tesis de Maestría, UAM-I, Abril **2009**.
79. Laskowski, M., Jr., & Finkenstadt, W. R. *Methods Enzymology*, **1972**, 26, 193–227.
80. Wyman, J. Linked Functions and Reciprocal Effects in Hemoglobin: A Second Look. *Advances in Protein Chemistry*, **1964**, 19, 223-286.
81. Turk, B.; Turk, V.; Turk, D. Structural and functional aspects of papain-like cysteine proteinases and their protein inhibitors. *Biological chemistry*, **1997**, 378, 141-150.
82. Gomes, M. T.; Teixeira, R. D.; Lopes, M. T.; Nagem, R. A.; Salas, C. E. X-ray crystal structure of CMS1MS2: a high proteolytic activity cysteine proteinase from *Carica candamarcensis*. *Amino Acids*, **2012**,43(6), 2381-91.
83. Gerhartz, B.; Engh, R.A.; Mentele, R.; Eckerskorn, C.; Torquato, R.; Wittmann, J.; Kolb, H.J.; Machleidt, W.; Fritz, H.; Auerswald, E. A. Quail cystatin: Isolation and characterisation of a new member of the cystatin family and its hypothetical interaction with cathepsin B. *FEBS Lett.*, **1997**, 412, 551-558.
84. Krause, J.; Bühner, M.; Sund, H. Studies of Glutamate Dehydrogenase. *Eur. J. Biochem.*, 1974, 41, 593 – 602.
85. Nordenman, B.; Björk, I. Binding of Low – Affinity and High – Affinity Heparin to Antitrombin. *Biochemistry*, 1978, 17(16), 3339 - 3344.
86. Kabsch, W.; Sander, C. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*, **1983**, 22(12), 2577-637.

## 9. APÉNDICES.

### A. Métodos desarrollados para purificar cistatina de pollo.

La purificación de cistatina de pollo (CEW) involucra elevados costos de operación debido a su complicado proceso de extracción y bajos rendimientos (alrededor de 20 a 28 % ) [16,29]. El precio comercial de CEW asciende a \$15700/mg (catalogo Sigma-Aldrich, 2014). Un objetivo inherente al presente trabajo fue implementar un proceso de extracción y purificación de CEW con un rendimiento aceptable, accesible a los recursos del laboratorio (materiales y equipos). Realizando modificaciones a los métodos reportados por otros grupos de investigación (Tabla A.1) se purifico CEW.

**Tabla A.1** Métodos implementados para la purificación de CEW.

Grupo de investigadores	Etapas de purificación	Y <sup>a</sup>
a) Anastasi <i>et al.</i> (1983) [29].	<ol style="list-style-type: none"> <li>1) Precipitación de ovomusina.</li> <li>2) Cromatografía de afinidad sobre Sefarosa–SCM-papaína.</li> <li>3) Cromatoenfoco.</li> </ol>	20 %
b) Turk <i>et al.</i> (1983) [48].	<ol style="list-style-type: none"> <li>1) Tratamiento alcalino (pH 12.0).</li> <li>2) Calentamiento a 62 °C.</li> <li>3) Cromatografía de exclusión molecular sobre Sephadex G-50.</li> <li>4) Cromatografía de intercambio iónico sobre DEAE-Sephacel.</li> </ol>	
c) Turk and Brzin (1990) [49].	<ol style="list-style-type: none"> <li>1) Calentamiento a 80° C.</li> <li>2) Cromatografía de afinidad sobre Sefarosa–SCM-papaína.</li> <li>1) Cromatografía de intercambio iónico en un sistema FPLC (Fast Protein Liquid Chromatography).</li> </ol>	32 %
d) Lindahl <i>et al.</i> (1988) [30].	<ol style="list-style-type: none"> <li>1) Precipitación de ovomusina.</li> <li>2) Cromatografía de afinidad sobre Sefarosa–SCM-papaína.</li> <li>3) Cromatografía de intercambio iónico sobre DEAE (grupos diethylaminoethyl con carga positiva).</li> </ol>	

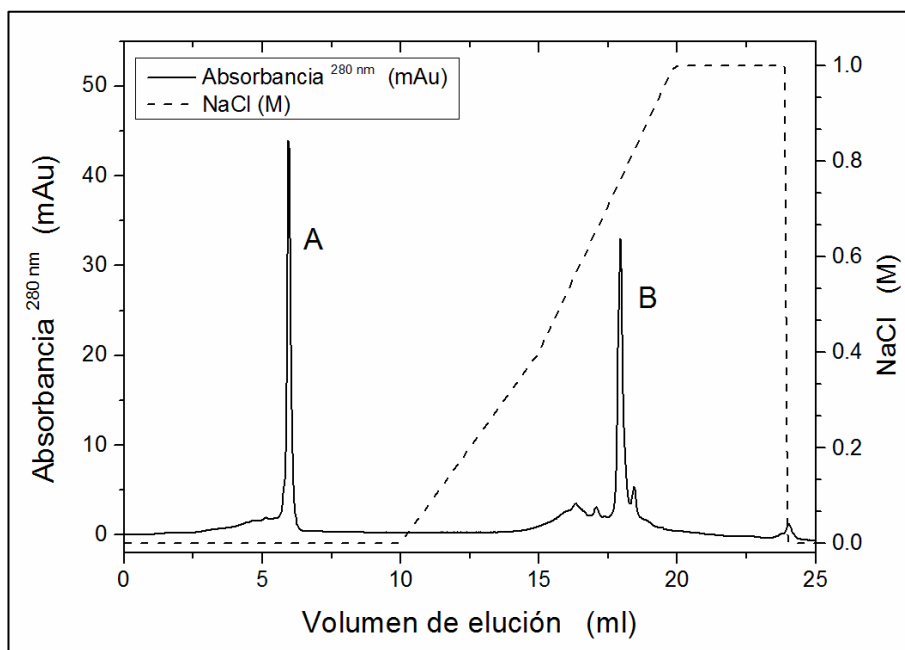
Continuación de Tabla A.1

e) Trziszka <i>et al.</i> (2004) [16].	<ol style="list-style-type: none"> <li>1) Tratamiento con pH (pH 3.0, pH 6.0).</li> <li>2) Cromatografía de afinidad sobre Sefarosa–SCM-papaína.</li> <li>3) Filtración en gel sobre Sephadex G-100.</li> <li>4) Cromatografía DEAE-Sephacel.</li> <li>5) Cromatografía en HPLC fase reversa sobre una columna de Nucleosil 100 C<sub>18</sub></li> </ol>	28 %
f) Trziszka <i>et al.</i> (2006) [50].	<ol style="list-style-type: none"> <li>1) Tratamiento con pH 4.0, alcohol 30 %.</li> <li>2) Cromatografía de afinidad sobre sefarosa–SCM-papaína.</li> <li>3) Filtración en gel sobre Sephadex G-100.</li> <li>4) Cromatografía DEAE-Sephacel.</li> <li>5) Cromatografía en HPLC fase reversa sobre una columna de Nucleosil 100 C<sub>18</sub>.</li> </ol>	
g) Wang and Wu (2014) [51].	<ol style="list-style-type: none"> <li>1) Tratamiento con pH 6.0.</li> <li>2) Cromatografía de intercambio catiónico en un sistema FPLC con una columna SP Sepharose FF. Cromatografía de afinidad sobre Sefarosa– SCM-papaína.</li> </ol>	33 %
h) Reyes-Espinosa (2015) desarrollado y presentado en el presente trabajo.	<ol style="list-style-type: none"> <li>1) Tratamiento con pH (pH 3.0, pH 6.0).</li> <li>2) Calentamiento a 80° C.</li> <li>3) Cromatografía de afinidad sobre Sefarosa– SCM-papaína.</li> <li>4) Cromatografía de exclusión molecular sobre una columna Sephadex 200 HiLoad 16/60 GP, seguida de una cromatografía en una columna Sephadex 200 10/300 GL, empleando un sistema FPLC.</li> </ol>	8.5 %

Donde: <sup>a</sup> Rendimiento reportado con respecto a la cantidad de cistatina de pollo (CEW) contenida en clara de huevo al inicio de la purificación (60 µg de cistatina / mL de clara [27]).

## B. Cromatografía de intercambio aniónico de CEW purificada.

La cromatografía de intercambio aniónico se realizó dializando una muestra de CEW purificada con amortiguador Tris-base 30 mM, pH 8.4. La muestra se cargó a una columna de intercambio aniónico Mono Q 5/50 GL (GE Healthcare) utilizando un gradiente bifásico de NaCl (0 a 0.4 M en 15 mL de  $V_e$  y de 0.4 a 0.8 M en 10 mL de  $V_e$ ).



**Figura B.1** Cromatograma de CEW purificada. Cromatografía de intercambio aniónico: columna Mono Q 5/50 GL, amortiguador de Tris-base 30 mM, pH 8.4.

La CEW presenta una composición de la fracción de cistatina de 0.33 (pico A), 0.3 (pico B) y el resto de contenido de proteína eluye en picos menores. Estos valores se calcularon con respecto al área total de los picos de elución con el software estadístico del equipo de FPLC y los datos del cromatograma presentado en la Figura B.1. El pico A corresponde a la forma no fosforilada ( $pI=6.5$ ) y el pico B ( $pI= 5.6$ ) a la fosforilada [31].



### C. Modelo para la determinación de la constante de disociación ( $K_d$ ) para un sistema proteína-ligando mediante el método de titulación fluorimétrica.

La reacción biomolecular reversible entre S-carboximetil-quimopapaína (SCM-quimopapaína) y cistatina de pollo (CEW) fue estudiada con el método de titulación fluorimétrica, el planteamiento de este modelo fue propuesto por Krause *et al.* (1974) [84].

La deducción de la ecuación del modelo para determinar la constante de disociación proteína–ligando en el equilibrio mediante el método titulación fluorimétrica que a continuación se describe fue presentada en la Tesis de Maestría de Reyes-Espinosa, 2009 [78]. La deducción parte de dos consideraciones [85]: La primera, el cambio en la cantidad física medida (Emisión de fluorescencia) es proporcional a la cantidad de complejo proteína ligando formado y la segunda, que todos los sitios de unión son equivalentes e independientes.

La reacción en el equilibrio entre SCM-quimopapaína y CEW está representada por ecuación C.1.



Donde  $P_{libre}$  es SCM-quimopapaína en forma libre,  $C_{libre}$  es CEW en forma libre y  $PC$  el complejo formado SCM-quimopapaína–CEW.

La constante de asociación en el equilibrio ( $K_a$ ) se define como:

$$K_a = \frac{1}{K_d} = \frac{[PC]}{[P_{libre}][C_{libre}]} \quad (C.2)$$

Donde:  $K_d$  es la constante de disociación en el equilibrio.

La cantidad total de SCM-quimopapaína ( $P$ ) contenida en la mezcla se pueden expresar como la cantidad en la forma libre ( $P_{libre}$ ) más la cantidad presente en forma de complejo ( $P_{PC}$ ):

$$P = P_{libre} + P_{PC} \quad (C.3)$$

La cantidad total de CEW ( $C$ ) contenidas en la mezcla se expresa como la cantidad en la forma libre ( $C_{libre}$ ) más la cantidad presente en forma de complejo ( $C_{PC}$ ):

$$C = C_{libre} + C_{PC} \quad (C.4)$$

La intensidad de fluorescencia emitida a una longitud de onda  $\lambda$  ( $F_\lambda$ ) depende de la intensidad de luz absorbida ( $I_A$ ) y del rendimiento cuántico ( $\phi_F$ ). Para absorbancias pequeñas a través de la ley de Beer y Lambert se puede demostrar:

$$F_\lambda = I_A \phi_F = 2.303 I_0 Z \phi_F \varepsilon l c \quad (\text{C.5})$$

Dónde:  $I_0$ , Intensidad de la luz de excitación;  $Z$ , factor instrumental, característico de cada instrumento;  $\varepsilon$ , coeficiente de extinción;  $l$ , longitud de la trayectoria óptica;  $c$ , concentración.

Definiendo  $k = 2.303 I_0 Z l$  y sustituyendo en ecuación C.5:

$$F_\lambda = k \phi_F \varepsilon c \quad (\text{C.6})$$

Dónde:  $k [=]$  cm;  $\varepsilon [=]$   $\text{cm}^{-1}\text{M}^{-1}$ ;  $c [=]$  M.

Definiendo la intensidad de fluorescencia de las especies ( $F_i$ ) contenidas en la celda de titulación como:

$F_{P \text{ libre}}$ , Fluorescencia de SCM-quimopapaína libre;

$F_P$ , Fluorescencia de SCM-quimopapaína total;

$F_{C \text{ libre}}$ , Fluorescencia de CEW libre;

$F_C$ , Fluorescencia de CEW total;

$F_M$ , Fluorescencia de la mezcla;

$F_{PC}$ , Fluorescencia del complejo SCM-quimopapaína–CEW.

De la ecuación C.3, la fluorescencia de SCM-quimopapaína contenida en la celda de reacción se puede expresar como:

$$F_P = F_{P \text{ libre}} + F_{PC} \quad (\text{C.7})$$

Para la fluorescencia de CEW contenida en la celda de reacción, de la ecuación C.4:

$$F_C = F_{C \text{ libre}} + F_{PC} \quad (\text{C.8})$$

La intensidad de fluorescencia de la mezcla se expresa como:

$$F_M = F_{P \text{ libre}} + F_{C \text{ libre}} + F_{PC} \quad (\text{C.9})$$

Desarrollando la ecuación C.7 en términos de C.6, se obtiene:

$$F_M = k\phi_P\varepsilon_P[P_{libre}] + k\phi_C\varepsilon_C[C_{libre}] + k\phi_{PC}\varepsilon_{PC}[PC] \quad (C.10)$$

Despejando  $P_{libre}$  de ecuación C.3,  $C_{libre}$  de ecuación C.4 y sustituyendo estas en ecuación C.9 y arreglando la ecuación se obtiene:

$$F_M = F_P + F_C + [PC](\phi_{PC}\varepsilon_{PC} - \phi_P\varepsilon_P - \phi_C\varepsilon_C) \quad (C.11)$$

$$\text{Dónde: } F_P = \phi_P\varepsilon_P[P] \quad (C.12)$$

$$F_C = \phi_C\varepsilon_C[C] \quad (C.13)$$

Arreglando ecuación C.11 y dividiendo esta por ecuación C.12, se obtiene:

$$Y = \frac{F_P + F_C - F_M}{F_P} = \frac{[PC]a}{[P]} \quad (C.14)$$

$$\text{Donde: } a = \frac{(\phi_{PC}\varepsilon_{PC} - \phi_P\varepsilon_P - \phi_C\varepsilon_C)}{\phi_P\varepsilon_P} \quad (C.15)$$

Despejando la concentración del complejo  $[PC]$  de C.14:

$$[PC] = \frac{Y[P]}{a} \quad (C.16)$$

Despejando  $C_{libre}$  de ecuación C.2 y sustituyendo  $P_{libre}$  obtenida de la ecuación C.3:

$$C_{libre} = \frac{[PC]K_d}{[P_{libre}]} = \frac{[PC]K_d}{[P] - [PC]} = \frac{K_d}{\frac{[P]}{[PC]} - 1} \quad (C.17)$$

Despejando  $\frac{[P]}{[PC]}$  de ecuación C.16 y sustituyendo en C.17:

$$C_f = \frac{K_d}{\frac{a}{Y} - 1} \quad (C.18)$$

Sustituyendo ecuación C.16 y ecuación C.18 en ecuación C.4:

$$C = \frac{K_d}{\frac{a}{Y} - 1} + \frac{Y[P]}{a} \quad (C.19)$$

Arreglando la ecuación (C.19) como:

$$\frac{[C]}{[P]} = \frac{Y K_d}{(a - Y)P_T} + \frac{Y}{a} \quad (C.20)$$

Definiendo:  $x = \frac{[C]}{[P]}$ ,  $w = \frac{K_d}{[P]}$ ; sustituyendo estas expresiones en la ecuación C.20:

$$x = \frac{wY}{(a - Y)} + \frac{Y}{a} \quad (C.21)$$

Desarrollando el polinomio de la ecuación C.21:

$$Y^2 - a(1 + w + x)Y + a^2x = 0 \quad (C.22)$$

La ecuación C.22 es un polinomio de segundo grado con coeficientes constantes, el cual se resuelve mediante la solución general:

$$Y = \frac{a}{2} \left( (1 + w + x) \pm \sqrt{(1 + w + x)^2 - 4x} \right) \quad (C.23)$$

Tomando la raíz negativa de la solución de la ecuación C.23:

$$Y = \frac{a}{2} \left( (1 + w + x) - \sqrt{(1 + w + x)^2 - 4x} \right) \quad (C.24)$$

Sustituyendo:  $x = \frac{[C]}{[P]}$ ,  $w = \frac{K_d}{[P]}$ , en la ecuación (C.24):

$$Y = \frac{a}{2} \left\{ \left( 1 + \frac{K_d}{[P]} + \frac{[C]}{[P]} \right) - \sqrt{\left[ \left( 1 + \frac{K_d}{[P]} + \frac{[C]}{[P]} \right)^2 - 4 \frac{[C]}{[P]} \right]} \right\} \quad (C.25)$$

Igualando ecuación C.14 y ecuación C.25 se obtiene el modelo matemático (ecuación C.26) con el cual se estudió la reacción biomolecular reversible entre SCM-quimopapaína y CEW:

$$Y = \frac{F_P + F_C - F_M}{F_P} = \frac{a}{2} \left\{ \left( 1 + \frac{K_d}{[P]} + \frac{[C]}{[P]} \right) - \sqrt{\left[ \left( 1 + \frac{K_d}{[P]} + \frac{[C]}{[P]} \right)^2 - 4 \frac{[C]}{[P]} \right]} \right\} \quad (C.26)$$

Donde;  $F_P$ ,  $F_C$ ,  $F_M$  son las variables experimentales (lecturas de emisión de fluorescencia);  $[C]$ ,  $[P]$  son variables independientes;  $K_d$ ,  $a$  son parámetros calculados a partir del ajuste no lineal de los datos experimentales evaluados en la primera parte de la ecuación sobre la curva  $Y$  vs.  $\frac{[C]}{[P]}$  (Segunda parte de la ecuación).

## D. Derivación del modelo modificado de transferencia de protón vinculado a la unión: *MPLM* (Crnogorac *et al.*, 2001), desarrollado para tres grupos isoacídicos.

La derivación que a continuación se presenta, es tomada del trabajo de Crnogorac *et al.* (2001) [34], la cual es una extensión con ligeras modificaciones de los trabajos de Laskowski and Finkenstadt (1972) [79] y Wyman (1964) [80]. La formulación desarrollada por Crnogorac *et al.* [34] para describir la dependencia de la constante de unión con respecto al cambio de pH fue empleado sin realizar modificación alguna para describir la dependencia del sistema en estudio. A continuación se presenta la derivación del modelo con la finalidad de presentar el aspecto conceptual de su construcción y formulación.

Dada la reacción de asociación entre la proteína P (Proteasa, quimopapaína) y la proteína I (Inhibidor, cistatina):

$$P + I \rightleftharpoons PI \quad (D.1)$$

Donde cada especie  $P, I, PI$  puede unir  $H^+$  (protones). La constante de unión ( $K$ ) se define como:

$$K = \frac{[PI]}{[P][I]} \quad (D.2)$$

Si la especie  $P$  posee  $n_P$  grupos titulables (sitios de unión de  $H^+$ ),  $P$  puede existir en  $2^{n_P}$  estados protonados. El número de estados protonados de  $P$  está dado por la ecuación D.3, donde  $i$  de los  $n_P$  grupos en  $P$  son desprotonados.

$$p_i^P = \binom{n_P}{n_P-i} = \frac{n_P!}{(n_P-i)!i!} \quad (D.3)$$

Estos  $p_i^P$  estados protonados son indexados con  $j$ . La concentración total de cada especie  $P, I$  y  $PI$  es una suma sobre todos sus posibles estados protonados y están dados por ecuación D.4, D.5 y D.6, respectivamente indexado con  $i$  y  $j$ .

$$[P] = \sum_{i=0}^{n_P} \sum_{j=1}^{p_i^P} [PH_i^j] \quad (D.4)$$

$$[I] = \sum_{i=0}^{n_I} \sum_{j=1}^{p_i^I} [IH_i^j] \quad (D.5)$$

$$[PI] = \sum_{i=0}^{n_{PI}} \sum_{j=1}^{p_i^{PI}} [PIH_i^j] \quad (D.6)$$

Sustituyendo D.4, D.5 y D.6 en D.2, se obtiene:

$$K = \frac{\sum_{i=0}^{n_{PI}} \sum_{j=1}^{p_i^{PI}} [PIH_i^j]}{\sum_{i=0}^{n_P} \sum_{j=1}^{p_i^P} [PH_i^j] \sum_{i=0}^{n_I} \sum_{j=1}^{p_i^I} [IH_i^j]} \quad (D.7)$$

Donde  $n_P, n_I$  y  $n_{PI}$  son los números de grupos titulables de  $P, I$  y  $PI$ , respectivamente.

El equilibrio entre las proteínas  $P$  y  $I$  completamente protonadas está dado por D.8, asumiendo que  $n_P + n_I = n_{PI}$ , es decir, que los sitios de unión de protones ( $H^+$ ) en la forma libre son los mismos en la forma asociada (no hay generación o creación de otros sitios).



La constante de equilibrio para la asociación de las proteínas completamente protonadas está definida como:

$$K_0 = \frac{[PI_0]}{[P_0][I_0]} \quad (D.9)$$

El equilibrio entre el estado completamente protonado de referencia y el estado que es indexado con  $i$  y  $j$ , para cada especie es expresado por D.10, D.12 y D.14, su respectiva constante de equilibrio D.11, D.13 y D.15.



$$L_0^{jP} = 1; \quad L_i^{jP} = \frac{[PH_i^j][H^+]^i}{[PH_0]} \quad (D.11)$$



$$L_0^{jI} = 1; \quad L_i^{jI} = \frac{[IH_i^j][H^+]^i}{[IH_0]} \quad (D.13)$$



$$L_0^{jI} = 1; \quad L_i^{jPI} = \frac{[PIH_i^j][H^+]^i}{[PIH_0]} \quad (D.15)$$

Dividiendo ecuación D.7 entre D.9 y ordenando, se obtiene:

$$K = K_0 \frac{\frac{\sum_{i=0}^{n_{PI}} \sum_{j=1}^{p_i^{PI}} [PIH_i^j]}{[PIH_0]}}{\left( \frac{\sum_{i=0}^{n_P} \sum_{j=1}^{p_i^P} [PH_i^j]}{[PH_0]} \right) \left( \frac{\sum_{i=0}^{n_I} \sum_{j=1}^{p_i^I} [IH_i^j]}{[IH_0]} \right)} \quad (D.16)$$

Utilizando ecuaciones D.11, D.13, D.15 y sustituyendo en la ecuación D.16:

$$K = K_0 \frac{\sum_{i=0}^{n_{PI}} \sum_{j=1}^{p_i^{PI}} [H^+]^{-i} L_i^{jPI}}{\left( \sum_{i=0}^{n_P} \sum_{j=1}^{p_i^P} [H^+]^{-i} L_i^{jP} \right) \left( \sum_{i=0}^{n_I} \sum_{j=1}^{p_i^I} [H^+]^{-i} L_i^{jI} \right)} \quad (D.17)$$

Aplicando log en ambos lados de la ecuación D.17 y diferenciando la ecuación resultante con respecto a  $\text{pH} = -\log[H^+]$ , se obtiene:

$$\frac{d\log K}{d\text{pH}} = \frac{d\log(K_0)}{d\text{pH}} + \frac{d\log\left(\sum_{i=0}^{n_{PI}} \sum_{j=1}^{p_i^{PI}} [H^+]^{-i} L_i^{jPI}\right)}{d\text{pH}} - \frac{d\log\left(\sum_{i=0}^{n_P} \sum_{j=1}^{p_i^P} [H^+]^{-i} L_i^{jP}\right)}{d\text{pH}} - \frac{d\log\left(\sum_{i=0}^{n_I} \sum_{j=1}^{p_i^I} [H^+]^{-i} L_i^{jI}\right)}{d\text{pH}} \quad (\text{D.18})$$

De la derivación de los últimos tres términos se obtiene:

$$\frac{d\log\left(\sum_{i=0}^{n_{PI}} \sum_{j=1}^{p_i^{PI}} [H^+]^{-i} L_i^{jPI}\right)}{d\text{pH}} = \left( \frac{-[H^+]}{\sum_{i=0}^{n_{PI}} \sum_{j=1}^{p_i^{PI}} [H^+]^{-i} L_i^{jPI}} \right) \frac{d\left(\sum_{i=0}^{n_{PI}} \sum_{j=1}^{p_i^{PI}} [H^+]^{-i} L_i^{jPI}\right)}{d[H^+]} = \frac{[H^+] \sum_{i=0}^{n_{PI}} \sum_{j=1}^{p_i^{PI}} i [H^+]^{-i-1} L_i^{jPI}}{\sum_{i=0}^{n_{PI}} \sum_{j=1}^{p_i^{PI}} [H^+]^{-i} L_i^{jPI}} = \langle i_{PI} \rangle \quad (\text{D.19})$$

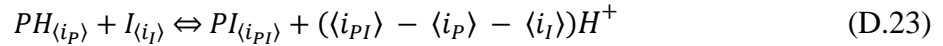
$$\frac{d\log\left(\sum_{i=0}^{n_P} \sum_{j=1}^{p_i^P} [H^+]^{-i} L_i^{jP}\right)}{d\text{pH}} = \left( \frac{-[H^+]}{\sum_{i=0}^{n_P} \sum_{j=1}^{p_i^P} [H^+]^{-i} L_i^{jP}} \right) \frac{d\left(\sum_{i=0}^{n_P} \sum_{j=1}^{p_i^P} [H^+]^{-i} L_i^{jP}\right)}{d[H^+]} = \frac{[H^+] \sum_{i=0}^{n_P} \sum_{j=1}^{p_i^P} i [H^+]^{-i-1} L_i^{jP}}{\sum_{i=0}^{n_P} \sum_{j=1}^{p_i^P} [H^+]^{-i} L_i^{jP}} = \langle i_P \rangle \quad (\text{D.20})$$

$$\frac{d\log\left(\sum_{i=0}^{n_I} \sum_{j=1}^{p_i^I} [H^+]^{-i} L_i^{jI}\right)}{d\text{pH}} = \left( \frac{-[H^+]}{\sum_{i=0}^{n_I} \sum_{j=1}^{p_i^I} [H^+]^{-i} L_i^{jI}} \right) \frac{d\left(\sum_{i=0}^{n_I} \sum_{j=1}^{p_i^I} [H^+]^{-i} L_i^{jI}\right)}{d[H^+]} = \frac{[H^+] \sum_{i=0}^{n_I} \sum_{j=1}^{p_i^I} i [H^+]^{-i-1} L_i^{jI}}{\sum_{i=0}^{n_I} \sum_{j=1}^{p_i^I} [H^+]^{-i} L_i^{jI}} = \langle i_I \rangle \quad (\text{D.21})$$

Sustituyendo D.19, D.20 y D.21 en D.18 se obtiene:

$$\frac{d\log K}{d\text{pH}} = q = \langle i_{PI} \rangle - \langle i_P \rangle - \langle i_I \rangle \quad (\text{D.22})$$

Donde  $\langle i_{PI} \rangle$ ,  $\langle i_P \rangle$  y  $\langle i_I \rangle$  representan el número promedio de  $H^+$  cedidos a partir de las formas de  $P$ ,  $I$  y  $PI$  completamente protonadas a el  $\text{pH}$  de interés. Por lo tanto,  $q$  es el número de  $H^+$  liberados bajo la asociación. El equilibrio a un  $\text{pH}$  determinado está definido por:



Integrando la ecuación D.22 sobre  $\text{pH}$ , se obtiene:

$$d\log K(\text{pH}_2) - d\log K(\text{pH}_1) = \int_{\text{pH}_1}^{\text{pH}_2} q \, d\text{pH} \quad (\text{D.24})$$

Por lo tanto, si se conoce la constante a  $\text{pH}_1$ , la constante a  $\text{pH}_2$  se puede conocer a través de la ecuación D.24, o bien, si la dependencia de  $K$  con respecto a la variación de  $\text{pH}$  es conocida, el número de protones liberados bajo la asociación está dada por la primera derivada de la ecuación D.24.

Para simplificar la ecuación D.17, Crnogorac *et al.* [34] definieron  $[P + I] = [P][I]$ , y empleando  $c$  como índice para la proteína asociada o combinada y  $s$  como índice para la proteína separada o libre. Si  $n$  es el número de  $H^+$  que pueden unirse a  $P$  y  $I$  ó  $PI$  ( $n = n_A + n_B = n_{AB}$ ) la ecuación D.17 se reescribe como:

$$K = K_0 \frac{\sum_{i=0}^n \sum_{j=1}^{p_i} [H^+]^{-i} L_i^{jc}}{\sum_{i=0}^n \sum_{j=1}^{p_i} [H^+]^{-i} L_i^{js}} \quad (D.25)$$

Donde  $L_i^{jc}$  y  $L_i^{js}$  están dados por:

$$L_0^{jc} = 1; \quad L_i^{jc} = \frac{[(P+I)H_i^{jc}][H^+]^i}{[(P+I)H_n]} \quad (D.26)$$

$$L_0^{js} = 1; \quad L_i^{js} = \frac{[(P+I)H_i^{js}][H^+]^i}{[(P+I)H_n]} \quad (D.27)$$

Para resolver D.25 se requieren conocer gran cantidad de parámetros, tomando como ejemplo el sistema de estudio cisteinproteasa–cistatina, se identificaron trece residuos ionizables (ver Tabla 4.6) posiblemente involucrados en la dependencia de la constante de asociación con respecto a la variación de  $pH$ , lo que implica considerar 16383 parámetros libres (valores de  $L_i^j = (2 \times 2^{13}) - 2 + \text{el valor de } K_0$ ).

Crnogorac *et al.* [34] modificaron este modelo introduciendo el concepto de grupos isoacídicos logrando reducir considerablemente el número de parámetros. Con el propósito de reducir el número de parámetros libres, uno puede asumir que un compuesto de  $pK_a$  puede ser asignado a residuos con comportamiento similar en  $pK_{aS}$ , estos residuos son llamados *grupos isoacídicos*.

Un grupo isoacídicos se define como como un grupo de residuos que titulan en el mismo intervalo de  $pH$ , sin requerir la interacción fuertes entre ellos [34]. Asumiendo que hay  $M$  grupos de residuos isoacídicos en la proteína. Cada grupo consiste de  $m_i$  residuos titulables. Reescribiendo D.25 se obtiene:

$$K = K_0 \frac{\sum_{i=0}^M \sum_{j=1}^{p_i} [H^+]^{-a_i^j} L_i^{jc}}{\sum_{i=0}^M \sum_{j=1}^{p_i} [H^+]^{-a_i^j} L_i^{js}} \quad (D.28)$$

$$a_i^j = \sum_{k=1}^M x_{i_k}^j m_k \quad (D.29)$$

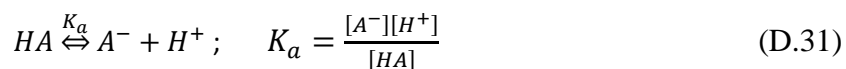
$$p_i = \binom{M}{M-i} = \frac{M!}{(M-1)!i!} \quad (D.30)$$

Donde  $p_i$  es el número de estados protonados teniendo  $i$  grupos de residuos isoacídicos desprotonados,  $i$  es el número de grupos de residuos isoacídicos desprotonados,  $j$  es el índice



de estado protonado,  $a_i^j$  es el número de protones disociados a partir de la proteína,  $x_{ik}^j$  es el  $k$ -ésimo elemento del estado protonado  $M$ -dimensional en el estado  $(i, j)$  comparado a su estado completamente protonado, dependiendo si  $i$  es protonado o desprotonado su vector tomara el valor de 1 o 0. Bajo este contexto un grupo de residuos isoacídicos es considerado totalmente protonado o totalmente desprotonado.

La constante de equilibrio de la disociación acida ( $K_a$ ) de un grupo titulable se define como:



La curva de titulación de una proteína con  $N$  residuos titulables puede ser escrita como la suma de  $N$  curvas de titulación estándar. Asumiendo que estos son  $M$  grupos de residuos titulables y cada grupo puede unir  $m_i$  protones, el equilibrio entre estados  $(i, j)$  y el completo estado protonado puede ser descrito como ecuación D.32 y D.33 para las proteínas asociadas y libres respectivamente (por definición,  $pK_a = -\log K_a$ ).

$$L_i^{js} = \prod_{k=1}^M x_{ik}^j (K_{a_k}^s)^{m_k} = 10^{-\sum_{k=1}^M x_{ik}^j m_k pK_{a_k}^s} \quad (D.32)$$

$$L_i^{jc} = \prod_{k=1}^M x_{ik}^j (K_{a_k}^c)^{m_k} = 10^{-\sum_{k=1}^M x_{ik}^j m_k pK_{a_k}^c} \quad (D.33)$$

Escribiendo la ecuación D.28 en términos de D.32 y D.33, y definiendo  $K_b$  como la constante de unión se obtiene:

$$K_b = K_0 \frac{\sum_{i=0}^M \sum_{j=1}^{p_i} [10^{-pH}]^{-a_i^j} 10^{-\sum_{k=1}^M x_{ik}^j m_k pK_{a_k}^c}}{\sum_{i=0}^M \sum_{j=1}^{p_i} [10^{-pH}]^{-a_i^j} 10^{-\sum_{k=1}^M x_{ik}^j m_k pK_{a_k}^s}} \quad (D.33)$$

A partir de la dependencia de pH de la asociación de SCM-quimopapaína–CEW (Figura 4.19), se pueden apreciar tres puntos de inflexión en la isoterma obtenida lo cual se puede relacionar con tres grupos de residuos isoacídicos ( $M=3$ ) que están cambiando su estado de protonación bajo la asociación. Por lo tanto se tienen  $2^3=8$  estados de protonación (D.34) para las proteínas en forma separadas, así como también, para las forma combinada.

Los correspondientes estados de protonación (ecuación D.30) para  $M=3$  están dados por:

$$p_i = \binom{M}{M-i} = \frac{M!}{(M-i)!i!}; \quad i = 0, 1, 2, 3 .$$

$$p_0 = \frac{3!}{(3)!0!} = 1; \quad p_1 = \frac{3!}{(2)!1!} = 3; \quad p_2 = \frac{3!}{(1)!2!} = 3; \quad p_3 = \frac{3!}{(0)!3!} = 1. \quad (D.34)$$

Los vectores de las expresiones indicadas en D.33 correspondientes a los estados de protonación  $(x_i^j)$  son:

$$\begin{aligned}
x_0^1 &= [0 \ 0 \ 0] \text{ (estado totalmente protonado, } x_{01}^1 = 0, x_{02}^1 = 0, x_{03}^1 = 0); \\
x_1^1 &= [1 \ 0 \ 0] \text{ (} x_{11}^1 = 1, x_{12}^1 = 0, x_{13}^1 = 0); \\
x_1^2 &= [0 \ 1 \ 0] \text{ (} x_{11}^2 = 0, x_{12}^2 = 1, x_{13}^2 = 0); \\
x_1^3 &= [0 \ 0 \ 1] \text{ (} x_{11}^3 = 0, x_{12}^3 = 0, x_{13}^3 = 1); \\
x_2^1 &= [1 \ 1 \ 0] \text{ (} x_{21}^1 = 1, x_{22}^1 = 1, x_{23}^1 = 0); \\
x_2^2 &= [1 \ 0 \ 1] \text{ (} x_{21}^2 = 1, x_{22}^2 = 0, x_{23}^2 = 1); \\
x_2^3 &= [0 \ 1 \ 1] \text{ (} x_{21}^3 = 0, x_{22}^3 = 1, x_{23}^3 = 1); \\
x_3^1 &= [1 \ 1 \ 1] \text{ (estado totalmente desprotonado, } x_{31}^1 = 1, x_{32}^1 = 1, x_{33}^1 = 1)
\end{aligned} \tag{D.35}$$

Desarrollando ecuación D.29 y sustituyendo los correspondientes valores de  $x_i^j$  (valores expresiones en D.35), se obtiene:

$$a_i^j = \sum_{k=1}^M x_{ik}^j m_k ; j = 1, 2, 3$$

$$a_0^1 = \sum_{k=1}^3 x_{0k}^1 m_k = x_{01}^1 m_1 + x_{02}^1 m_2 + x_{03}^1 m_3 = 0 \tag{D.36}$$

$$a_1^1 = \sum_{k=1}^3 x_{1k}^1 m_k = x_{11}^1 m_1 + x_{12}^1 m_2 + x_{13}^1 m_3 = m_1 \tag{D.37}$$

$$a_1^2 = \sum_{k=1}^3 x_{1k}^2 m_k = x_{11}^2 m_1 + x_{12}^2 m_2 + x_{13}^2 m_3 = m_2 \tag{D.38}$$

$$a_1^3 = \sum_{k=1}^3 x_{1k}^3 m_k = x_{11}^3 m_1 + x_{12}^3 m_2 + x_{13}^3 m_3 = m_3 \tag{D.39}$$

$$a_2^1 = \sum_{k=1}^3 x_{2k}^1 m_k = x_{21}^1 m_1 + x_{22}^1 m_2 + x_{23}^1 m_3 = m_1 + m_2 \tag{D.40}$$

$$a_2^2 = \sum_{k=1}^3 x_{2k}^2 m_k = x_{21}^2 m_1 + x_{22}^2 m_2 + x_{23}^2 m_3 = m_1 + m_3 \tag{D.41}$$

$$a_2^3 = \sum_{k=1}^3 x_{2k}^3 m_k = x_{21}^3 m_1 + x_{22}^3 m_2 + x_{23}^3 m_3 = m_2 + m_3 \tag{D.42}$$

$$a_3^1 = \sum_{k=1}^3 x_{3k}^1 m_k = x_{31}^1 m_1 + x_{32}^1 m_2 + x_{33}^1 m_3 = m_1 + m_2 + m_3 \tag{D.43}$$

Desarrollando para la forma separada de la proteína, la ecuación D.32 para  $M=3$ .

$$L_0^{js} = 10^{-\sum_{k=1}^3 x_{0k}^j m_k pK_{a_k}^s} = 10^{-(x_{01}^j m_1 pK_{a_1}^s + x_{02}^j m_2 pK_{a_2}^s + x_{03}^j m_3 pK_{a_3}^s)} \tag{D.44}$$

$$L_1^{js} = 10^{-\sum_{k=1}^3 x_{1k}^j m_k pK_{a_k}^s} = 10^{-(x_{11}^j m_1 pK_{a_1}^s + x_{12}^j m_2 pK_{a_2}^s + x_{13}^j m_3 pK_{a_3}^s)} \tag{D.45}$$

$$L_2^{js} = 10^{-\sum_{k=1}^3 x_{2k}^j m_k pK_{a_k}^s} = 10^{-(x_{21}^j m_1 pK_{a_1}^s + x_{22}^j m_2 pK_{a_2}^s + x_{23}^j m_3 pK_{a_3}^s)} \tag{D.46}$$

$$L_3^{js} = 10^{-\sum_{k=1}^3 x_{3k}^j m_k pK_{a_k}^s} = 10^{-(x_{31}^j m_1 pK_{a_1}^s + x_{32}^j m_2 pK_{a_2}^s + x_{33}^j m_3 pK_{a_3}^s)} \quad (D.47)$$

Desarrollando para la forma combinada de la proteína, la ecuación D.33 para  $M=3$ .

$$L_0^{jc} = 10^{-\sum_{k=1}^3 x_{0k}^j m_k pK_{a_k}^c} = 10^{-(x_{01}^j m_1 pK_{a_1}^c + x_{02}^j m_2 pK_{a_2}^c + x_{03}^j m_3 pK_{a_3}^c)} \quad (D.48)$$

$$L_1^{jc} = 10^{-\sum_{k=1}^3 x_{1k}^j m_k pK_{a_k}^c} = 10^{-(x_{11}^j m_1 pK_{a_1}^c + x_{12}^j m_2 pK_{a_2}^c + x_{13}^j m_3 pK_{a_3}^c)} \quad (D.49)$$

$$L_2^{jc} = 10^{-\sum_{k=1}^3 x_{2k}^j m_k pK_{a_k}^c} = 10^{-(x_{21}^j m_1 pK_{a_1}^c + x_{22}^j m_2 pK_{a_2}^c + x_{23}^j m_3 pK_{a_3}^c)} \quad (D.50)$$

$$L_3^{jc} = 10^{-\sum_{k=1}^3 x_{3k}^j m_k pK_{a_k}^c} = 10^{-(x_{31}^j m_1 pK_{a_1}^c + x_{32}^j m_2 pK_{a_2}^c + x_{33}^j m_3 pK_{a_3}^c)} \quad (D.51)$$

Sustituyendo los correspondientes valores de  $x_i^j$  (expresiones indicadas en D.35) en ecuaciones D.44 a D.51, se obtiene:

a) Para las proteínas en forma separada (libre):

$$\begin{aligned} L_0^{1s} &= 10^{-(x_{01}^1 m_1 pK_{a_1}^s + x_{02}^1 m_2 pK_{a_2}^s + x_{03}^1 m_3 pK_{a_3}^s)} = \\ &= 10^{-((0) m_1 pK_{a_1}^s + (0) m_2 pK_{a_2}^s + (0) m_3 pK_{a_3}^s)} = 10^0 = 1 \end{aligned} \quad (D.52)$$

$$\begin{aligned} L_1^{1s} &= 10^{-(x_{11}^1 m_1 pK_{a_1}^s + x_{12}^1 m_2 pK_{a_2}^s + x_{13}^1 m_3 pK_{a_3}^s)} = \\ &= 10^{-((1) m_1 pK_{a_1}^s + (0) m_2 pK_{a_2}^s + (0) m_3 pK_{a_3}^s)} = 10^{-m_1 pK_{a_1}^s} \end{aligned} \quad (D.53)$$

$$\begin{aligned} L_2^{1s} &= 10^{-(x_{21}^1 m_1 pK_{a_1}^s + x_{22}^1 m_2 pK_{a_2}^s + x_{23}^1 m_3 pK_{a_3}^s)} = \\ &= 10^{-((0) m_1 pK_{a_1}^s + (1) m_2 pK_{a_2}^s + (0) m_3 pK_{a_3}^s)} = 10^{-m_2 pK_{a_2}^s} \end{aligned} \quad (D.54)$$

$$\begin{aligned} L_3^{1s} &= 10^{-(x_{31}^1 m_1 pK_{a_1}^s + x_{32}^1 m_2 pK_{a_2}^s + x_{33}^1 m_3 pK_{a_3}^s)} = \\ &= 10^{-((0) m_1 pK_{a_1}^s + (0) m_2 pK_{a_2}^s + (1) m_3 pK_{a_3}^s)} = 10^{-m_3 pK_{a_3}^s} \end{aligned} \quad (D.55)$$

$$\begin{aligned} L_2^{1s} &= 10^{-(x_{21}^1 m_1 pK_{a_1}^s + x_{22}^1 m_2 pK_{a_2}^s + x_{23}^1 m_3 pK_{a_3}^s)} = \\ &= 10^{-((1) m_1 pK_{a_1}^s + (1) m_2 pK_{a_2}^s + (0) m_3 pK_{a_3}^s)} = 10^{-(m_1 pK_{a_1}^s + m_2 pK_{a_2}^s)} \end{aligned} \quad (D.56)$$

$$\begin{aligned} L_2^{2s} &= 10^{-(x_{21}^2 m_1 pK_{a_1}^s + x_{22}^2 m_2 pK_{a_2}^s + x_{23}^2 m_3 pK_{a_3}^s)} = \\ &= 10^{-((1) m_1 pK_{a_1}^s + (0) m_2 pK_{a_2}^s + (1) m_3 pK_{a_3}^s)} = 10^{-(m_1 pK_{a_1}^s + m_3 pK_{a_3}^s)} \end{aligned} \quad (D.57)$$

$$\begin{aligned} L_2^{3s} &= 10^{-(x_{21}^3 m_1 pK_{a_1}^s + x_{22}^3 m_2 pK_{a_2}^s + x_{23}^3 m_3 pK_{a_3}^s)} = \\ &= 10^{-((0) m_1 pK_{a_1}^s + (1) m_2 pK_{a_2}^s + (1) m_3 pK_{a_3}^s)} = 10^{-(m_2 pK_{a_2}^s + m_3 pK_{a_3}^s)} \end{aligned} \quad (D.58)$$

$$\begin{aligned} L_3^{1s} &= 10^{-(x_{31}^1 m_1 pK_{a_1}^s + x_{32}^1 m_2 pK_{a_2}^s + x_{33}^1 m_3 pK_{a_3}^s)} = \\ &= 10^{-((1) m_1 pK_{a_1}^s + (1) m_2 pK_{a_2}^s + (1) m_3 pK_{a_3}^s)} = 10^{-((m_1 pK_{a_1}^s + m_2 pK_{a_2}^s + m_3 pK_{a_3}^s))} \end{aligned} \quad (D.59)$$

b) Para las proteínas en forma combinada (asociada):

$$L_0^{1c} = 10^{-(x_{01}^1 m_1 pK_{a1}^c + x_{02}^1 m_2 pK_{a2}^c + x_{03}^1 m_3 pK_{a3}^c)} = 10^{-((0) m_1 pK_{a1}^c + (0) m_2 pK_{a2}^c + (0) m_3 pK_{a3}^c)} = 10^0 = 1 \quad (D.60)$$

$$L_1^{1c} = 10^{-(x_{11}^1 m_1 pK_{a1}^c + x_{12}^1 m_2 pK_{a2}^c + x_{13}^1 m_3 pK_{a3}^c)} = 10^{-((1) m_1 pK_{a1}^c + (0) m_2 pK_{a2}^c + (0) m_3 pK_{a3}^c)} = 10^{-m_1 pK_{a1}^c} \quad (D.61)$$

$$L_1^{2c} = 10^{-(x_{11}^2 m_1 pK_{a1}^c + x_{12}^2 m_2 pK_{a2}^c + x_{13}^2 m_3 pK_{a3}^c)} = 10^{-((0) m_1 pK_{a1}^c + (1) m_2 pK_{a2}^c + (0) m_3 pK_{a3}^c)} = 10^{-m_2 pK_{a2}^c} \quad (D.62)$$

$$L_1^{3c} = 10^{-(x_{11}^3 m_1 pK_{a1}^c + x_{12}^3 m_2 pK_{a2}^c + x_{13}^3 m_3 pK_{a3}^c)} = 10^{-((0) m_1 pK_{a1}^c + (0) m_2 pK_{a2}^c + (1) m_3 pK_{a3}^c)} = 10^{-m_3 pK_{a3}^c} \quad (D.63)$$

$$L_2^{1c} = 10^{-(x_{21}^1 m_1 pK_{a1}^c + x_{22}^1 m_2 pK_{a2}^c + x_{23}^1 m_3 pK_{a3}^c)} = 10^{-((1) m_1 pK_{a1}^c + (1) m_2 pK_{a2}^c + (0) m_3 pK_{a3}^c)} = 10^{-(m_1 pK_{a1}^c + m_2 pK_{a2}^c)} \quad (D.64)$$

$$L_2^{2c} = 10^{-(x_{21}^2 m_1 pK_{a1}^c + x_{22}^2 m_2 pK_{a2}^c + x_{23}^2 m_3 pK_{a3}^c)} = 10^{-((1) m_1 pK_{a1}^c + (0) m_2 pK_{a2}^c + (1) m_3 pK_{a3}^c)} = 10^{-(m_1 pK_{a1}^c + m_3 pK_{a3}^c)} \quad (D.65)$$

$$L_2^{3c} = 10^{-(x_{21}^3 m_1 pK_{a1}^c + x_{22}^3 m_2 pK_{a2}^c + x_{23}^3 m_3 pK_{a3}^c)} = 10^{-((0) m_1 pK_{a1}^c + (1) m_2 pK_{a2}^c + (1) m_3 pK_{a3}^c)} = 10^{-(m_2 pK_{a2}^c + m_3 pK_{a3}^c)} \quad (D.66)$$

$$L_3^{1c} = 10^{-(x_{31}^1 m_1 pK_{a1}^c + x_{32}^1 m_2 pK_{a2}^c + x_{33}^1 m_3 pK_{a3}^c)} = 10^{-((1) m_1 pK_{a1}^c + (1) m_2 pK_{a2}^c + (1) m_3 pK_{a3}^c)} = 10^{-((m_1 pK_{a1}^c + m_2 pK_{a2}^c + m_3 pK_{a3}^c))} \quad (D.67)$$

Rescribiendo ecuación D.28 y desarrollando para  $M=3$ , definiendo:  $K_0 = K_{\text{int}}$ , constante intrínseca:

$$K = K_0 \frac{\sum_{i=0}^M \sum_{j=1}^{p_i} [H^+]^{-a_i^j} L_i^{jc}}{\sum_{i=0}^M \sum_{j=1}^{p_i} [H^+]^{-a_i^j} L_i^{js}} = K_{\text{int}} \frac{[Num]}{[Denom]} \quad (D.68)$$

Donde:

$$\begin{aligned} Num = & \sum_{i=0}^M \sum_{j=1}^{p_i} [H^+]^{-a_i^j} L_i^{jc} = \sum_{j=1}^{p_0} [H^+]^{-a_0^j} L_0^{jc} + \sum_{j=1}^{p_1} [H^+]^{-a_1^j} L_1^{jc} + \\ & \sum_{j=1}^{p_2} [H^+]^{-a_2^j} L_2^{jc} + \sum_{j=1}^{p_3} [H^+]^{-a_3^j} L_3^{jc} = [H^+]^{-a_0^1} L_0^{1c} + [H^+]^{-a_1^1} L_1^{1c} + \\ & [H^+]^{-a_1^2} L_1^{2c} + [H^+]^{-a_1^3} L_1^{3c} + [H^+]^{-a_2^1} L_2^{1c} + [H^+]^{-a_2^2} L_2^{2c} + \\ & [H^+]^{-a_2^3} L_2^{3c} + [H^+]^{-a_3^1} L_3^{1c} \end{aligned}$$

$$\begin{aligned}
Deno &= \sum_{i=0}^M \sum_{j=1}^{p_i} [H^+]^{-a_i^j} L_i^{j_s} = \sum_{j=1}^{p_0} [H^+]^{-a_0^j} L_0^{j_s} + \sum_{j=1}^{p_1} [H^+]^{-a_1^j} L_1^{j_s} + \\
&\quad \sum_{j=1}^{p_2} [H^+]^{-a_2^j} L_2^{j_s} + \sum_{j=1}^{p_3} [H^+]^{-a_3^j} L_3^{j_s} = [H^+]^{-a_0^1} L_0^{1_s} + [H^+]^{-a_1^1} L_1^{1_s} + \\
&\quad [H^+]^{-a_1^2} L_1^{2_s} + [H^+]^{-a_1^3} L_1^{3_s} + [H^+]^{-a_2^1} L_2^{1_s} + [H^+]^{-a_2^2} L_2^{2_s} + \\
&\quad [H^+]^{-a_2^3} L_2^{3_s} + [H^+]^{-a_3^1} L_3^{1_s}
\end{aligned}$$

Sustituyendo expresiones D.52 a D.67 en D.58, finalmente se obtiene la ecuación del modelo *MPLM*, para tres grupos isoacídicos (ecuación D.69) que permite describir la dependencia de  $K_b$  con respecto a la variación de pH.

$$K_b(\text{pH}) = K_{\text{int}}[\text{Num}/\text{Deno}] \quad (\text{D.69})$$

Donde: *Num* y *Deno* están definidos por:

$$\begin{aligned}
Num &= 1 + 10^{-m_1 \text{pKa}_1^c} \times 10^{m_1 \text{pH}} + 10^{-m_2 \text{pKa}_2^c} \times 10^{m_2 \text{pH}} + 10^{-m_3 \text{pKa}_3^c} \times 10^{m_3 \text{pH}} + \\
&\quad 10^{-(m_1 \text{pKa}_1^c + m_2 \text{pKa}_2^c)} \times 10^{(m_1 + m_2) \text{pH}} + 10^{-(m_1 \text{pKa}_1^c + m_3 \text{pKa}_3^c)} \times 10^{(m_1 + m_3) \text{pH}} + \\
&\quad 10^{-(m_2 \text{pKa}_2^c + m_3 \text{pKa}_3^c)} \times 10^{(m_2 + m_3) \text{pH}} + 10^{-(m_1 \text{pKa}_1^c + m_2 \text{pKa}_2^c + m_3 \text{pKa}_3^c)} \times \\
&\quad 10^{(m_1 + m_2 + m_3) \text{pH}}
\end{aligned}$$

$$\begin{aligned}
Deno &= 1 + 10^{-m_1 \text{pKa}_1^s} \times 10^{m_1 \text{pH}} + 10^{-m_2 \text{pKa}_2^s} \times 10^{m_2 \text{pH}} + 10^{-m_3 \text{pKa}_3^s} \times 10^{m_3 \text{pH}} + \\
&\quad 10^{-(m_1 \text{pKa}_1^s + m_2 \text{pKa}_2^s)} \times 10^{(m_1 + m_2) \text{pH}} + 10^{-(m_1 \text{pKa}_1^s + m_3 \text{pKa}_3^s)} \times 10^{(m_1 + m_3) \text{pH}} + \\
&\quad 10^{-(m_2 \text{pKa}_2^s + m_3 \text{pKa}_3^s)} \times 10^{(m_2 + m_3) \text{pH}} + 10^{-(m_1 \text{pKa}_1^s + m_2 \text{pKa}_2^s + m_3 \text{pKa}_3^s)} \times \\
&\quad 10^{(m_1 + m_2 + m_3) \text{pH}}
\end{aligned}$$

Los parámetros libres a estimar en esta ecuación (D.69) son diez: la constante intrínseca ( $K_{\text{int}}$ ) tres grupos isoacídicos ( $m_1$ ,  $m_2$ ,  $m_3$ , cada uno constituidos de tres posibles residuos), tres valores de  $\text{pK}_a$  para la forma asociada ( $\text{pK}_{a_1}^c$ ,  $\text{pK}_{a_2}^c$  y  $\text{pK}_{a_3}^c$ ) y tres valores de  $\text{pK}_a$  para la forma libre ( $\text{pK}_{a_1}^s$ ,  $\text{pK}_{a_2}^s$  y  $\text{pK}_{a_3}^s$ ).

## E. Propiedades estructurales quimopapaína, CEW y proteínas de complejos homólogos cisteinproteasa–cistatina.

El alineamiento de secuencia primaria de estructuras homólogas cisteinproteasa cistatina (Sección 3.4.3) es presentado en la Tabla E.1. La composición de aminoácidos y las propiedades estructurales de las proteínas que componen estos heterodímeros son presentados en la Tabla E.2 y E.3, respectivamente.

**Tabla E.1** Alineamiento de secuencia primaria<sup>1</sup> de estructuras homólogas cisteinproteasa–cistatina (Tabla 4.9) e identificación de residuos localizados en la interfase<sup>2</sup>.

---

Alineamiento de secuencia de cadena correspondiente a cistatina (Inhibidor).

```

---GAPVPVDENDEGLQRALQFAMAEYNRASNDKYSSRVVRVISAKRQLVSGIKYILQVE 57
---GAPVPVDENDEGLQRALQFAMAEYNRASNDKYSSRVVRVISAKRQLVSGIKYILQVE 57
MMSGAPSATQPATAETQHIADQVRSQLEEKYNKKFP--VFKAVSFKSQVVAGTNYFIKVVH 58
MIPGGLSEAKPATPEIQEIVDKVKPQLEEKTNETYG--KLEAVQYKTQVVAGTNYIYIKVR 58
MIPGGLSEAKPATPEIQEIVDKVKPQLEEKTNETYG--KLEAVQYKTQVVAGTNYIYIKVR 58
MIPGGLSEAKPATPEIQEIVDKVKPQLEEKTNETYG--KLEAVQYKTQVVAGTNYIYIKVR 58
MIPGGLSEAKPATPEIQEIVDKVKPQLEEKTNETYG--KLEAVQYKTQVVAGTNYIYIKVR 58
      * .   ..      * .   : . . : : . * . . :   . . . . . * * : * : * : * : * .

```

---

```

IGRTTCPKSSGDLQSCEFHDPEPEMAKYTTCTFVVYSIPWLNQIKLLESKCQ 108 (Quimo–CEW)
IGRTTCPKSSGDLQSCEFHDPEPEMAKYTTCTFVVYSIPWLNQIKLLESKCQ 108 (Pap–CEW)
VGDEDF-----VHLRQVFSLP--HENKPLTLSNYQTNKAKHDELTYF--- 98 (1STF)
AGDNKY-----MHLKVFKSLP--GQNEIDLVLTYGYQVDKNKDDELTYF--- 98 (3K9M)
AGDNKY-----MHLKVFKSLP--GQNEIDLVLTYGYQVDKNKDDELTYF--- 98 (1NB5)
AGDNKY-----MHLKVFKSLP--GQNEIDLVLTYGYQVDKNKDDELTYF--- 98 (3KSE)
AGDNKY-----MHLKVFKSLP--GQNEIDLVLTYGYQVDKNKDDELTYF--- 98 (3KFQ)
      *           : :      * : . *      :           : :      * .           : :      : *

```

---

Donde: <sup>1</sup> El alineamiento se realizó con el servidor EMBL-EBI, Tools > Multiple Sequence Alignment> ClustalW2: <http://www.ebi.ac.uk/Tools/msa/clustalw2/>. <sup>2</sup> El análisis de los residuos que constituyen la interfase se realizó con el servidor PDBePISA (Proteins, Interfaces, Structures and Assemblies): <http://www.ebi.ac.uk/pdbe/pisa/>. Los residuos (código de una letra) sombreados son los residuos identificados en la interfase. Los residuos marcados en recuadro son los residuos ionizables identificados en la interfase del heterodímero señalados en la Tabla 4.9.

## Continuación Tabla E.1.

## Alineamiento de secuencia de cadena correspondiente a cisteinproteasa (enzima).

YPQSIDWRAKG---- <td>54</td>	54	
IPEYVDWRQKG---- <td>54</td>	54	
IPEYVDWRQKG---- <td>54</td>	54	
LPASFDAREQWPQCPTIKEIRDQGCSCWAFGAVEAISDRICHTNAHVSVEVSAEDLL	60	
YPPSMDWRKKN---FVSPVKNQGCSCWTFSTTGALES AVAIATGKMLSLAE--QQLV	55	
APRSVDWREKG---- <td>54</td>	54	
LPKSVDWRKKN---YVTPVKNQKQGCSCWAFSATGALEGQMFRTGKLVSLSE--QNLV	54	
* . * * : . : : * * * * * : : : . * . . : * * :		
DCDKH--SYGCKGGYQTTSLQYVANN--GVHTSKVYPYQAKQYKCR-----	96	
DCDRR--SYGCNGGYPWSALQLVAQY--GIHYRNTYPYEGVQRYCR-----	96	
DCDRR--SYGCNGGYPWSALQLVAQY--GIHYRNTYPYEGVQRYCR-----	96	
TCCGSMCGDGCNGGYPAEAWNFWRKGLVSGGLYESHVGCRPYSIPPCEHHVNGSRPPCT	120	
DCAQNFNNHGCQGLPSQAFEYIRYNGKIMGEDTYPYKGQDDHCK-----	100	
DCSGPQNEGCNGGLMDYAFQYVQDNGGLDSEESYPYEATEESCK-----	99	
DCSRPQNGQGCNGGFMARAFQYVKENGGLDSEESYPYVAVDEICK-----	99	
* . * * : * : : : : . : . .		
-----ATDKPGPKVKITGYKRVPSNCETSFLGALAN-QPLSVLV EAGGK P	140	
-----SREKGPYAAKTGVRQVQPYNQALLYSIAN-QPVSVVLQAAGK D	140	
-----SREKGPYAAKTGVRQVQPYNQALLYSIAN-QPVSVVLQAAGK D	140	
GEGDTPKCSKICEPGYSPTYKQDKHYGNSYSVSNSEKDIMAEIYKNGPVEGAFSVYSD-	179	
-----FQPDKA-IAFVKDVANITMNDDEAMVEAVALYNPVSFAFEVTND-	143	
-----YNPKYS-VANDAGF-VDI PKQEKALMKAVATVGPI SV AIDAGH E S	142	
-----YRPENS-VAQDTGFTV VAPGKEKALMKAVATVGPI SVAMDAGHSS	143	
: : : * : . . . .		
FQLYKSGVFDG-PCGT--KLDH AVTAVGYGTS----DGKNYII IKNSWGPNWGEKGYMR	192	
FQLYRGGIFVG-PCGN--KVDH AVAAVGYGP-----NYILIKNSWGTGWGENGYIR	188	
FQLYRGGIFVG-PCGN--KVDH AVAAVGYGP-----NYILIKNSWGTGWGENGYIR	188	
FLLYKSGVYQHVTGEM--MGGH AIRILGWGVEN----GTPYWL VANSWNT DWGDN GFFK	232	
FLMYRKG IYSSTSCHKTPDKVNH AVLAVGYGEE----NGIPYWIVKNSWGPQWGMNGYFL	199	
FLFYKEGIYFEPDCSS--EDMDHGVLVVGYGFESTESDDNKYWLVKNSWGE EWGMGGYVK	200	
FQFYKSGIYFEPDCSS--KNLDHGVLVVGYGFEGANS DNSKYWLVKNSWGP EWG SNGYVK	201	
* : * : * : : . * : : * * * * * : * : * * * * * : .		
LKRQSGNSQGTGCVYKSSYYPFKGFA	218	Quimopapaína-CEW (Quimo-CEW)
IKRGTGNSYGVCGLYTSSFPVKN--	212	Papaína-CEW (Pap-CEW)
IKRGTGNSYGVCGLYTSSFPVKN--	212	Papaína-Estefina B (1STF)
ILRGQD---HCGIES E VVAGIPRTD	254	Catepsina B-Estefina A (3K9M)
IERGK---NMCGLAACASYPIPLV-	220	Catepsina H-Estefina A (1NB5)
MAKDRR---NHCGIASAASYPTV---	220	Catepsina L-Estefina A (3KSE)
IAKDKN---NHCGIATAASYPNV---	221	Catepsina V-Estefina A (3KFQ)
: * * :		

**Tabla E.2** Composición de residuos de aminoácidos<sup>a</sup> de proteínas que componen las estructuras homólogas cisteinproteasa–cistatina en comparación (Tabla E.1).

Proteínas	Quimo.	Papaína	Cat. B	Cat. H	Cat. L	Cat. V	CEW	CA	CB
<b>Polar ácido</b>									
Asp (D)	7	6	13	10	14	10	5	6	4
Glu (E)	7	7	17	11	18	11	8	9	6
<b>Polar básico</b>									
Lys (K)	21	10	10	14	13	19	7	12	9
Arg (R)	5	12	9	4	6	5	6	1	2
His (H)	3	2	9	4	3	3	1	1	5
<b>Polar no cargado</b>									
Gly (G)	27	28	31	20	25	24	5	8	3
Ser (S)	17	13	22	14	19	20	10	2	7
Thr (T)	14	8	11	9	7	7	5	7	7
Cys (C)	8	6	14	8	6	6	4	0	0
Tyr (Y)	15	19	13	12	13	11	5	6	4
Asn (N)	11	13	11	15	10	14	4	5	5
Gln (Q)	11	13	6	9	8	9	7	6	7
<b>No polar</b>									
Ala (A)	14	14	14	19	17	18	7	5	8
Val (V)	16	18	16	15	16	20	9	9	9
Leu (L)	12	11	10	11	10	10	8	8	6
Ile (I)	8	12	14	10	6	5	6	4	2
Pro (P)	11	10	15	12	9	10	5	5	5
Phe (F)	6	4	7	10	8	9	3	2	7
Met (M)	1	0	4	8	6	4	2	2	2
Trp (W)	4	5	8	5	5	5	1	0	0
<b>Total de residuos</b>	<b>218</b>	<b>211</b>	<b>254</b>	<b>220</b>	<b>219</b>	<b>220</b>	<b>108</b>	<b>98</b>	<b>98</b>
<b>Masa molecular<sup>B</sup> (KDa)</b>	<b>23.66</b>	<b>23.32</b>	<b>27.82</b>	<b>24.31</b>	<b>24.04</b>	<b>23.91</b>	<b>12.19</b>	<b>11.01</b>	<b>11.16</b>

**Donde:** <sup>a</sup> Obtenidos a partir de los respectivos PDBs (citados en Sección 3.4.3); <sup>B</sup>, masa molecular calculada en base a la composición de aminoácidos, empleando el servidor PROTEIN CALCULATOR v3.4 (<http://protcalc.sourceforge.net/>); Quimo., quimopapaína; Cat., catepsina; CA, estefina A; CB, estefina B.



**Tabla E.3** Propiedades estructurales de las proteínas que componen los heterodímeros indicados en Tabla E.1.

Cadena	% HSP <sup>a</sup>	% Hélice $\alpha$ <sup>b</sup>	% Hoja $\beta$ <sup>b</sup>	Cistinas	Cisteínas
Quimopapaína		24	23	3 C <sup>22</sup> -C <sup>63</sup> , C <sup>56</sup> -C <sup>95</sup> , C <sup>153</sup> -C <sup>204</sup>	C <sup>25*</sup> y C <sup>117</sup>
Papaína	57.6 <sup>1</sup>	26	21	3 C <sup>22</sup> -C <sup>63</sup> , C <sup>56</sup> -C <sup>95</sup> , C <sup>153</sup> -C <sup>200</sup>	C <sup>25*</sup>
Catepsina B	23.1 <sup>1</sup>	25	25	6 C <sup>14</sup> -C <sup>43</sup> , C <sup>26</sup> -C <sup>71</sup> , C <sup>62</sup> -C <sup>128</sup> , C <sup>63</sup> -C <sup>71</sup> , C <sup>108</sup> -C <sup>119</sup> , C <sup>100</sup> -C <sup>132</sup>	C <sup>29*</sup> y C <sup>240</sup>
Catepsina H	40.0 <sup>1</sup>	24	22	3 C <sup>22</sup> -C <sup>63</sup> , C <sup>56</sup> -C <sup>95</sup> , C <sup>154</sup> -C <sup>200</sup>	C <sup>25*</sup> y C <sup>205</sup>
Catepsina L	41.0 <sup>1</sup>	28	21	3 C <sup>22</sup> -C <sup>65</sup> , C <sup>56</sup> -C <sup>98</sup> , C <sup>156</sup> -C <sup>209</sup>	C <sup>25*</sup>
Catepsina V	42.7 <sup>1</sup>	32	18	3 C <sup>22</sup> -C <sup>65</sup> , C <sup>56</sup> -C <sup>98</sup> , C <sup>157</sup> -C <sup>210</sup>	C <sup>25*</sup>
CEW		15	38	2 C <sup>71</sup> -C <sup>81</sup> , C <sup>95</sup> -C <sup>115</sup>	
Estefina A		18	41	–	
Estefina B	53.1 <sup>2</sup>	19	39	–	

Donde: <sup>a</sup> Resultado de alineamiento (<sup>1</sup> respecto a quimopapaína y <sup>2</sup> a estefina A; Tabla E.1);  
<sup>b</sup> valor de % de estructura secundaria asignada con el programa DSSP [86], disponible en repositorio Protein Data Bank (PDB) [20]. \*, Residuo catalítico; Identificación de cistinas usando el servidor firstglance (<http://firstglance.jmol.org>).

## **F. Residuos *hot spots* en la interface de complejos homólogos cisteinproteasa–cistatina: Computational alanine scanning of protein-protein.**

Kortemme *et al.* [72,73] desarrollaron una herramienta computacional rápida para la predicación de residuos de aminoácidos energéticamente importantes localizados en la interface proteína–proteína, al cual nombraron “*computational alanine scanning*”, en el presente trabajo el método lo voy a citar como *CAS* por sus siglas. Los datos de entrada son el archivo PDB de un complejo proteína–proteína; retornando al término del proceso de datos una lista de residuos *hot spots*, o cadenas laterales de aminoácidos que se predice desestabilizan significativamente la interfase cuando son hipotéticamente mutados a residuo de alanina, análogos a resultados experimentales de mutagénesis *alanine-scanning*. La metodología de cálculo (presentada en referencia 81) fue validada con el exitoso diseño de interfaces de proteínas con nueva especificidad y actividad. También ha aportado nuevos conocimientos sobre los mecanismos de la especificidad del receptor y reacciones secundarias en los sistemas biológicos [73].

Si método *CAS* se aplica a la estructura de un complejo sin datos experimentales existentes (como ocurre en para la interacción quimopapaína CEW) el algoritmo de computo identifica todos los residuos de la interfase de proteína–proteína. En su metodología, un residuo localizado en la interfase lo definen como aquel residuo que al menos uno de sus átomos se encuentra a 4 Å de cualquier átomo perteneciente a la cadena opuesta, o bien, un residuo que se queda enterrado (oculto) significativamente en la formación del complejo, lo cual se mide por un aumento de átomos  $C_{\beta}$  dentro de una esfera con radio de 8 Å alrededor del átomo de  $C_{\beta}$  del residuo de interés. El programa entonces reemplaza cada residuo de la interface individualmente con residuos de alanina, y realiza el cómputo del efecto de esta mutación sobre la energía libre de unión del complejo [73].

El análisis con el método *CAS* permite identificar residuos neutros y residuos *hot spots*. Los cuales son definidos de acuerdo al cambio en la energía libre de unión ( $\Delta\Delta G_{\text{bind}}$ ), ya sea menor o mayor a 1 kcal/mol respectivamente, cuando el residuo es reemplazado por alanina (para más detalles del método consultar referencia 72,73).

En la Tabla F.1, presento resultados del análisis computacional con el método CAS. Las estructuras analizadas fueron las citadas en Sección 3.4.3.

**Tabla F.1** Análisis computacional *alanine scanning*<sup>i</sup> de complejos homólogos cisteinproteasa–cistatina. (Parte correspondiente a cisteinproteasa, Enzima).

Quim-CEW	Pap-CEW	1STF	3K9M	1NB5	3KSE	3KFQ
Quimo.	Papaína	Papaína	Cat. B	Cat. H	Cat. L	Cat. V
	<b>N<sup>18</sup></b> <u>0.04</u> /1.00 <sup>(1)</sup>	<b>N<sup>18</sup></b> <u>0.01</u> /1.07 <sup>(1)</sup>	<b>D<sup>22</sup></b> <u>-0.08</u> /2.58 <sup>(1)</sup>		<b>N<sup>18</sup></b> <u>0.06</u> /1.23 <sup>(1)</sup>	<b>N<sup>18</sup></b> <u>0.07</u> /1.32 <sup>(1)</sup>
<b>Q<sup>19</sup></b> 1.94/2.46 <sup>(1)</sup>	<b>Q<sup>19</sup></b> 1.33/1.71 <sup>(1)</sup>	<b>Q<sup>19</sup></b> 1.45/2.71 <sup>(1)</sup>	<b>Q<sup>23</sup></b> 1.60/1.87 <sup>(1)</sup>	<b>Q<sup>20</sup></b> <u>0.44</u> /5.30 <sup>(0)</sup>	<b>Q<sup>19</sup></b> 1.58/2.46 <sup>(1)</sup>	<b>Q<sup>19</sup></b> 3.26/2.49 <sup>(1)</sup>
<b>C<sup>22</sup></b> <u>-0.01</u> /5.11 <sup>(0)</sup>	<b>C<sup>22</sup></b> <u>-0.02</u> /4.89 <sup>(0)</sup>	<b>C<sup>22</sup></b> <u>-0.03</u> /4.79 <sup>(0)</sup>	<b>C<sup>26</sup></b> <u>-0.03</u> /5.05 <sup>(0)</sup>	<b>C<sup>22</sup></b> <u>-0.01</u> /4.93 <sup>(0)</sup>	<b>Q<sup>21</sup></b> <u>0.96</u> /-0.98 <sup>(1)*</sup>	<b>C<sup>22</sup></b> <u>0.00</u> /4.84 <sup>(0)</sup>
<b>C<sup>25</sup></b> <u>0.00</u> /4.97 <sup>(0)</sup>	<b>C<sup>25</sup></b> <u>0.00</u> /5.20 <sup>(0)</sup>	<b>W<sup>26</sup></b> <u>0.33</u> /5.78 <sup>(1)</sup>	<b>S<sup>28</sup></b> <u>0.00</u> /1.27 <sup>(0)</sup>	<b>C<sup>26</sup></b> <u>-0.04</u> /5.08 <sup>(1)</sup>	<b>C<sup>22</sup></b> <u>-0.02</u> /4.92 <sup>(0)</sup>	<b>W<sup>26</sup></b> <u>0.35</u> /5.70 <sup>(1)</sup>
			<b>C<sup>29</sup></b> <u>0.03</u> /5.23 <sup>(1)</sup>	<b>W<sup>27</sup></b> <u>0.27</u> /5.32 <sup>(1)</sup>	<b>W<sup>26</sup></b> <u>0.36</u> /5.64 <sup>(1)</sup>	
			<b>C<sup>71</sup></b> <u>-0.02</u> /4.90 <sup>(0)</sup>			
<b>Y<sup>61</sup></b> <u>0.50</u> /1.37 <sup>(0)</sup>	<b>Y<sup>61</sup></b> <u>0.68</u> /1.02 <sup>(1)</sup>	<b>Y<sup>61</sup></b> <u>0.44</u> /1.46 <sup>(1)</sup>	<b>N<sup>72</sup></b> 3.72/ <u>0.74</u> <sup>(1)</sup>	<b>C<sup>66</sup></b> <u>-0.02</u> /4.85 <sup>(1)</sup>	<b>C<sup>65</sup></b> <u>-0.02</u> /4.97 <sup>(0)</sup>	<b>C<sup>65</sup></b> <u>-0.01</u> /5.00 <sup>(0)</sup>
<b>Y<sup>67</sup></b> <u>0.55</u> /1.22 <sup>(1)</sup>		<b>C<sup>63</sup></b> <u>-0.09</u> /5.02 <sup>(1)</sup>	<b>Y<sup>75</sup></b> 1.05/ <u>0.17</u> <sup>(1)</sup>	<b>Q<sup>67</sup></b> <u>7.03</u> / <u>-0.55</u> <sup>(1)</sup>		<b>F<sup>69</sup></b> 1.34/ <u>-0.19</u> <sup>(1)</sup>
		<b>Y<sup>67</sup></b> 1.64/ <u>0.98</u> <sup>(1)</sup>	<b>C<sup>108</sup></b> <u>-0.13</u> /5.13 <sup>(1)</sup>	<b>L<sup>70</sup></b> <u>0.35</u> /1.44 <sup>(1)</sup>		<b>M<sup>70</sup></b> <u>0.25</u> /1.95 <sup>(1)</sup>
			<b>E<sup>109</sup></b> 1.77/ <u>-0.05</u> <sup>(1)</sup>			
			<b>R<sup>116</sup></b> 1.07/ <u>0.04</u> <sup>(1)</sup>			
			<b>C<sup>119</sup></b> <u>-0.02</u> /3.99 <sup>(0)</sup>	<b>E<sup>139</sup></b> <u>-0.01</u> /2.13 <sup>(0)</sup>		
<b>F<sup>141</sup></b> <u>0.04</u> /3.13 <sup>(0)</sup>	<b>F<sup>141</sup></b> <u>0.06</u> /3.23 <sup>(0)</sup>	<b>V<sup>133</sup></b> <u>0.25</u> /1.87 <sup>(1)</sup>	<b>F<sup>174</sup></b> <u>0.04</u> /4.32 <sup>(0)</sup>	<b>V<sup>140</sup></b> <u>0.39</u> /1.83 <sup>(1)</sup>		
<b>Q<sup>142</sup></b> 1.78/1.26 <sup>(1)</sup>	<b>Q<sup>142</sup></b> 1.77/1.12 <sup>(1)</sup>	<b>Q<sup>142</sup></b> <u>0.51</u> /1.27 <sup>(1)</sup>	<b>V<sup>176</sup></b> <u>0.48</u> /2.20 <sup>(1)</sup>	<b>L<sup>145</sup></b> 1.01/1.25 <sup>(1)</sup>	<b>L<sup>144</sup></b> 1.18/ <u>0.98</u> <sup>(1)</sup>	
	<b>R<sup>145</sup></b> 2.80/ <u>0.72</u> <sup>(1)</sup>	<b>V<sup>157</sup></b> <u>0.13</u> /1.56 <sup>(0)</sup>	<b>L<sup>181</sup></b> <u>0.93</u> /1.58 <sup>(1)</sup>	<b>V<sup>164</sup></b> <u>0.12</u> /2.34 <sup>(0)</sup>	<b>M<sup>161</sup></b> <u>0.03</u> /1.99 <sup>(0)</sup>	<b>L<sup>162</sup></b> <u>0.02</u> /4.09 <sup>(0)</sup>
			<b>E<sup>194</sup></b> 1.50/ <u>-0.48</u> <sup>(1)</sup>	<b>N<sup>165</sup></b> <u>0.00</u> /2.47 <sup>(1)</sup>	<b>D<sup>162</sup></b> <u>-0.41</u> /1.25 <sup>(1)</sup>	<b>D<sup>163</sup></b> <u>-0.35</u> /1.48 <sup>(1)</sup>
<b>H<sup>159</sup></b> 1.51/1.99 <sup>(0)</sup>	<b>H<sup>159</sup></b> <u>0.43</u> /2.34 <sup>(1)</sup>	<b>H<sup>159</sup></b> 1.61/1.84 <sup>(0)</sup>	<b>M<sup>196</sup></b> <u>0.43</u> /1.13 <sup>(1)</sup>	<b>H<sup>166</sup></b> <u>0.71</u> /1.91 <sup>(1)</sup>	<b>H<sup>163</sup></b> <u>0.85</u> /1.49 <sup>(1)</sup>	<b>H<sup>164</sup></b> 1.47/1.82 <sup>(0)</sup>
<b>S<sup>180</sup></b> 1.62/2.07 <sup>(0)</sup>	<b>W<sup>177</sup></b> 4.05/3.22 <sup>(1)</sup>	<b>W<sup>177</sup></b> 3.10/4.42 <sup>(1)</sup>	<b>H<sup>199</sup></b> 1.11/1.36 <sup>(1)</sup>	<b>W<sup>188</sup></b> 2.37/3.33 <sup>(1)</sup>	<b>W<sup>189</sup></b> 3.23/4.22 <sup>(1)</sup>	<b>W<sup>190</sup></b> 3.02/3.85 <sup>(1)</sup>
<b>W<sup>181</sup></b> 2.99/2.79 <sup>(1)</sup>	<b>W<sup>181</sup></b> 0.55/5.44 <sup>(1)</sup>	<b>W<sup>181</sup></b> <u>0.55</u> /5.98 <sup>(1)</sup>	<b>W<sup>221</sup></b> 2.20/3.8 <sup>(1)</sup>	<b>W<sup>192</sup></b> <u>0.70</u> /5.37 <sup>(1)</sup>	<b>W<sup>193</sup></b> <u>0.46</u> /5.77 <sup>(0)</sup>	<b>W<sup>194</sup></b> <u>0.52</u> /5.41 <sup>(1)</sup>
<b>W<sup>185</sup></b> <u>0.59</u> /4.24 <sup>(1)</sup>		<b>F<sup>207</sup></b> <u>0.24</u> /2.08 <sup>(1)</sup>	<b>W<sup>225</sup></b> <u>0.60</u> /5.65 <sup>(1)</sup>			

Donde: <sup>i</sup> Disponible en el servidor robetta (<http://robetta.bakerlab.org/alaninescan>); Los archivos PDB corresponden a las siguientes estructuras: Quim-CEW (quimopapaína–cistatina de pollo), Pap-CEW (papaína–cistatina de pollo), 1STF (papaína–estefina B), 3K9M (catepsina B–estefina A), 1NB5 (catepsina H–estefina A), 3KSE (catepsina L–estefina A) y 3KFQ (catepsina V–estefina A); Cat, catepsina.

Continuación Tabla F.1. (Parte correspondiente a cistatina, Inhibidor).

Quim-CEW	Pap-CEW	1STF	3K9M	1NB5	3KSE	3KFQ
CEW	CEW	Estefina B	Estefina A	Estefina A	Estefina A	Estefina A
			<b>M<sup>1</sup></b> 1.38/0.11 <sup>(1)</sup>			
		<b>M<sup>2</sup></b> 1.05/0.10 <sup>(1)</sup>	<b>I<sup>2</sup></b> 1.11/0.18 <sup>(1)</sup>	<b>I<sup>2</sup></b> 1.42/0.04 <sup>(1)</sup>	<b>I<sup>2</sup></b> 1.46/-0.04 <sup>(1)</sup>	<b>I<sup>2</sup></b> 1.25/0.05 <sup>(1)</sup>
		<b>S<sup>7</sup></b> -0.01/2.44 <sup>(0)</sup>	<b>L<sup>6</sup></b> 0.14/2.46 <sup>(0)</sup>	<b>L<sup>6</sup></b> 0.25/1.86 <sup>(1)</sup>	<b>L<sup>6</sup></b> 0.25/2.61 <sup>(0)</sup>	<b>L<sup>6</sup></b> 0.15/2.68 <sup>(0)</sup>
			<b>S<sup>7</sup></b> 0.16/2.40 <sup>(1)</sup>		<b>S<sup>7</sup></b> 0.01/2.41 <sup>(0)</sup>	
			<b>Y<sup>43</sup></b> 0.97/3.87 <sup>(1)</sup>			
			<b>T<sup>45</sup></b> 0.73/1.45 <sup>(1)</sup>			
<b>Q<sup>53</sup></b> 1.27/0.06 <sup>(1)</sup>	<b>Q<sup>53</sup></b> 1.37/0.74 <sup>(1)</sup>	<b>Q<sup>46</sup></b> 2.47/4.31 <sup>(1)</sup>	<b>Q<sup>46</sup></b> 5.18/7.34 <sup>(1)</sup>	<b>Q<sup>46</sup></b> 6.21/7.13 <sup>(1)</sup>	<b>Q<sup>46</sup></b> 2.76/7.86 <sup>(1)</sup>	<b>Q<sup>46</sup></b> 2.80/7.64 <sup>(1)</sup>
<b>L<sup>54</sup></b> 1.51/0.35 <sup>(1)</sup>	<b>L<sup>54</sup></b> 2.08/0.58 <sup>(1)</sup>	<b>V<sup>47</sup></b> 1.4/0.48 <sup>(1)</sup>	<b>V<sup>47</sup></b> 1.82/0.48 <sup>(1)</sup>	<b>V<sup>47</sup></b> 1.49/0.54 <sup>(1)</sup>	<b>V<sup>47</sup></b> 1.13/0.29 <sup>(1)</sup>	<b>V<sup>47</sup></b> 1.18/0.38 <sup>(1)</sup>
			<b>V<sup>48</sup></b> 0.74/1.51 <sup>(1)</sup>	<b>V<sup>48</sup></b> 0.69/1.72 <sup>(1)</sup>	<b>V<sup>48</sup></b> 0.89/1.44 <sup>(1)</sup>	<b>V<sup>48</sup></b> 0.91/1.46 <sup>(1)</sup>
<b>V<sup>55</sup></b> 1.00/1.10 <sup>(1)</sup>	<b>V<sup>55</sup></b> 1.36/0.74 <sup>(1)</sup>	<b>T<sup>51</sup></b> 0.02/1.35 <sup>(1)</sup>	<b>T<sup>51</sup></b> 0.23/1.23 <sup>(1)</sup>	<b>T<sup>51</sup></b> -0.01/1.08 <sup>(1)</sup>	<b>T<sup>51</sup></b> 0.05/1.22 <sup>(1)</sup>	<b>T<sup>51</sup></b> 0.49/1.14 <sup>(1)</sup>
<b>I<sup>58</sup></b> 0.63/1.60 <sup>(1)</sup>	<b>I<sup>58</sup></b> 0.48/1.51 <sup>(1)</sup>	<b>N<sup>52</sup></b> 5.59/4.27 <sup>(1)</sup>	<b>N<sup>52</sup></b> 3.5/1.97 <sup>(1)</sup>	<b>N<sup>52</sup></b> 3.14/2.62 <sup>(1)</sup>	<b>N<sup>52</sup></b> 3.26/1.90 <sup>(1)</sup>	<b>N<sup>52</sup></b> 2.98/2.01 <sup>(1)</sup>
	<b>K<sup>59</sup></b> 1.66/1.48 <sup>(1)</sup>	<b>R<sup>68</sup></b> 3.06/4.14 <sup>(1)</sup>	<b>Y<sup>53</sup></b> 0.11/2.77 <sup>(1)</sup>			
			<b>Y<sup>54</sup></b> 0.12/7.74 <sup>(1)</sup>	<b>Y<sup>54</sup></b> 2.79/8.57 <sup>(1)</sup>		
<b>Y<sup>100</sup></b> 1.20/2.38 <sup>(1)</sup>	<b>Y<sup>100</sup></b> 1.74/2.80 <sup>(1)</sup>	<b>F<sup>70</sup></b> 0.60/2.18 <sup>(1)</sup>	<b>F<sup>70</sup></b> 1.16/2.39 <sup>(1)</sup>	<b>F<sup>70</sup></b> 0.78/2.12 <sup>(1)</sup>	<b>F<sup>70</sup></b> 1.28/2.61 <sup>(1)</sup>	<b>F<sup>70</sup></b> 1.08/2.51 <sup>(1)</sup>
<b>I<sup>102</sup></b> 0.05/1.54 <sup>(0)</sup>	<b>I<sup>102</sup></b> 0.05/1.79 <sup>(0)</sup>	<b>L<sup>73</sup></b> 1.96/1.06 <sup>(1)</sup>	<b>L<sup>73</sup></b> 2.10/0.80 <sup>(1)</sup>	<b>L<sup>73</sup></b> 2.07/1.10 <sup>(1)</sup>	<b>L<sup>73</sup></b> 1.99/0.85 <sup>(1)</sup>	<b>L<sup>73</sup></b> 1.87/1.03 <sup>(1)</sup>
<b>W<sup>104</sup></b> 3.87/1.02 <sup>(1)</sup>	<b>W<sup>104</sup></b> 3.86/0.77 <sup>(1)</sup>	<b>H<sup>75</sup></b> 1.34/-0.42 <sup>(1)</sup>	<b>Q<sup>76</sup></b> 3.62/-0.52 <sup>(1)</sup>	<b>Q<sup>76</sup></b> 1.59/-0.14 <sup>(1)</sup>		<b>Q<sup>76</sup></b> 1.01/0.13 <sup>(0)</sup>
	<b>L<sup>105</sup></b> 0.12/1.14 <sup>(0)</sup>		<b>T<sup>83</sup></b> 0.29/1.24 <sup>(0)</sup>		<b>T<sup>83</sup></b> 0.00/1.09 <sup>(1)</sup>	<b>T<sup>83</sup></b> -0.01/1.20 <sup>(1)</sup>
	<b>N<sup>106</sup></b> 1.30/0.27 <sup>(1)</sup>	<b>Y<sup>97</sup></b> 1.94/1.95 <sup>(1)</sup>	<b>F<sup>98</sup></b> -0.01/-1.17 <sup>(0)*</sup>	<b>F<sup>98</sup></b> 0.00/-1.25 <sup>(0)*</sup>		

Donde: En letras con formatos negritas señalo al residuo de aminoácido en código de una letra, el superíndice con el mismo formato indica la posición en la secuencia de estructura primaria de la proteína respectiva; El superíndice entre paréntesis indica el número de residuo localizado en la interfase, si un átomo del residuo se encontró a 4 Å de cualquier átomo perteneciente a la cadena opuesta se le asigna el valor de 1 y si no tiene contacto pero es oculto bajo la unión el valor es 0; los dos números en subíndice separados por una barra diagonal denotan: El primer número muestra el valor de la predicción del cambio en la energía libre de unión ( $\Delta\Delta G_{\text{bind}}^{\text{CAS}}$ ) bajo la hipotética mutación de alanina, el segundo predice el cambio en la estabilidad de la proteína sola ( $\Delta G_{\text{bind}}^{\text{CAS}}$ ) bajo la hipotética mutación de alanina y los números subrayados indican los cambio en energía libre (tanto para ( $\Delta\Delta G_{\text{bind}}^{\text{CAS}}$  como  $\Delta G_{\text{bind}}^{\text{CAS}}$ ) menores a 1 Kcal/mol, residuos no *hot spots*; Las celdas sombreadas indican un cambio mayor a 1 Kcal/mol en  $\Delta\Delta G_{\text{bind}}^{\text{CAS}}$ , estos residuos son nombrados residuos *hot spots*.

Los aspectos de la interacción la proteína–proteína que no se pueden predecir bien, debido a la simplificación de este modelo (citadas por los autores) [72,73], son las siguientes: (i) En muchos casos, a pesar de que el residuo pueda ser identificado como *hot spot*, la magnitud del efecto electrostático no puede ser capturada; particularmente notorio en Asp y Glu, este caso es más claro en tratamientos más precisos de la electrostática en cálculos en proteínas, necesarios para inducir efectos de polarización y corrimientos de las constantes de ionización. (ii) No se puede predecir los efectos de la mutación hipotética de residuos que forman enlaces de hidrógeno con moléculas de aguas localizadas en la interfase; los enlaces específicos con moléculas de agua no son considerados y los efectos de solvatación solo son tratados de manera implícita. (iii) Solo limitados cambios conformacionales en la cadena son tomados en consideración

# Effects of pH on the Association between the Inhibitor Cystatin and the Proteinase Chymopapain

Francisco Reyes-Espinosa, Alfonso Arroyo-Reyna\*, Ponciano García-Gutiérrez, Iris N. Serratos and Rafael A. Zubillaga\*

Departamento de Química, Universidad Autónoma Metropolitana-Iztapalapa. Apartado Postal 55-534, Iztapalapa, México D.F., C.P 09340, México



Rafael A. Zubillaga

**Abstract:** Cysteine proteinases are involved in many aspects of physiological regulation. In humans, some cathepsins have shown another function in addition to their role as lysosomal proteases in intracellular protein degradation; they have been implicated in the pathogenesis of several heart and blood vessel diseases and in cancer development. In this work, we present a fluorometric and computational study of the binding of one representative plant cysteine proteinase, chymopapain, to one of the most studied inhibitors of these proteinases: chicken cystatin. The binding equilibrium constant,  $K_b$ , was determined in the pH range between 3.5 and 10.0, revealing a maximum in the affinity at pH 9.0. We constructed an atomic model for the chymopapain–cystatin dimer by docking the individual 3D protein structures; subsequently, the model was refined using a 100 ns NPT molecular dynamics simulation in explicit water. Upon scrutiny of this model, we identified 14 ionizing residues at the interface of the complex using a cutoff distance of 5.0 Å. Using the  $pK_a$  values predicted with PROPKA and a modified proton-linkage model, we performed a regression analysis on our data to obtain the composite  $pK_a$  values for three isoacidic residues. We also calculated the electrostatic component of the binding energy ( $\Delta G_{b,elec}$ ) at different pH values using an implicit solvent model and APBS software. The pH profile of this calculated energy compares well with the experimentally obtained binding energy,  $\Delta G_b$ . We propose that the residues that form an interchain ionic pair, Lys139A from chymopapain and Glu19B from cystatin, as well as Tyr61A and Tyr67A from chymopapain are the main residues responsible for the observed pH dependence in the chymopapain–cystatin affinity.

**Keywords:** Cysteine proteinase, enzyme–inhibitor interactions, fluorometric titration, molecular dynamics simulation, pH dependence,  $pK_a$  shift, protein–protein docking.

## INTRODUCTION

The most well-characterized family of cysteine proteases (CPs) is the papain family, which includes various human cathepsins. In addition to their role as lysosomal proteases in intracellular protein degradation, cysteine cathepsins are also present in the extracellular space and have been reported to be involved in the pathogenesis of several heart and blood vessel diseases [1]. These proteins are also implicated in cancer development, and cathepsin inhibitors have been proposed as anticancer agents [2]. A family of inhibitors known as cystatins modulates the activity of CPs. Therefore, an improved understanding of the mechanisms through which these inhibitors interact with CPs could contribute to the development of novel inhibitors to prevent inappropriate proteolysis in cardiovascular diseases and cancer.

Cysteine proteinases from papaya latex exhibit molecular weights in the range 23.5–24.9 kDa [3] and differ in their isoelectric points (pI), which are all basic, in contrast to the

pI values ranging from 4.1–7.1 for animal cathepsins. Chymopapain, the most abundant cysteine proteinase in papaya latex, is present as isoforms with isoelectric points in the pH range 10.1–10.6 [4]. Comparing of the aligned amino acid sequences indicates that chymopapain has 58.0 % identity with papain and 39.9 % identity with human cathepsin L. In addition, the polypeptide chains of these proteins share a common folding pattern; when comparing the crystallographic structure of chymopapain (PDB code 1YAL) with those of papain (PDB code 9PAP) and human cathepsin L (PDB code 2Y2J) using the positional root mean square deviation (RMSD) of the backbones, we obtained values of 1.0 Å and 1.2 Å, respectively.

One of the most studied protein inhibitors of CP has been chicken cystatin. This inhibitor is a 12.5 kDa protein from chicken egg white (CEW-cystatin) that forms equimolar and reversible complexes with papain-like cysteine proteinases and can also bind to these proteins when the reactive cysteine has been blocked [5]. The three-dimensional structure of CEW-cystatin has been determined, and a docking model for the interaction of the inhibitor and papain has been proposed [6]. CEW-cystatin occurs in two forms that have isoelectric points of 6.5 and 5.6, one of these forms is phosphorylated at Ser80 [7]; both forms, phosphorylated and unphosphory-

\*Address correspondence to these author at the Departamento de Química, Universidad Autónoma Metropolitana-Iztapalapa. Apartado Postal 55-534, Iztapalapa, México D.F., C.P 09340, México; Tel: (55) 5804 4674; Fax: (55) 5804 4666; E-mails: [aar@xanum.uam.mx](mailto:aar@xanum.uam.mx); [zlr@xanum.uam.mx](mailto:zlr@xanum.uam.mx)

lated, exhibit very similar affinities for active papain [8] and for the tight-binding inactivated forms of the enzyme [9].

Protein-protein affinities often exhibit a pH dependence. A previous study demonstrated that the affinity between CEW-cystatin and s-(carboxymethyl)-papain decreases with pH in the 8.2–10.5 range [9]. This result prompted us to investigate the pH dependence of the interaction between CEW-cystatin and s-(carboxymethyl)-chymopapain. In this work, we determined the binding equilibrium constant,  $K_b$ , for this reaction in the pH range between 3.5 and 10.0 and observed a maximum in the affinity at pH 9.0. We also present a modeled structure of the CEW-cystatin–chymopapain dimer that was obtained by docking the individual atomic structures and refining the docked structure using a molecular dynamics simulation in explicit water. This model allowed us to identify the interfacial ionizing residues and to evaluate their contribution to the binding electrostatic energy. Analyzing the experimental binding data using a modified proton-linkage model based on the modeled dimer and electrostatic analyses allowed us to propose which residues could be mainly responsible for the observed pH dependence of the affinity.

## MATERIALS AND METHODS

### Proteins

Chicken cystatin was purified from chicken egg white using a procedure modified from previous reports [5, 8, 10–12]. CEW-cystatin was obtained from a homogenate of egg white diluted with an equal volume of 0.1 M NaCl. The homogenate was brought to pH 3.0 using 3.0 M HCl and incubated for 1 h at 4°C. The precipitated protein was removed by centrifugation at 4000 g for 10 min; then, the supernatant pH was adjusted to 6.0 using 3.0 M NaOH, and the supernatant was incubated for 1 h at room temperature. The solution was heated at 80°C for 3 min and cooled in a water bath. The precipitated ovomucin was removed by centrifugation at 4000 g for 20 min. The supernatant was subjected to affinity chromatography by pouring the solution into a gel of s-(carboxymethyl)-papain–sepharose 4B (SCM-papain–sepharose) equilibrated with 50 mM phosphate buffer, pH 7.0. The proteins bound to SCM-papain–sepharose were desorbed using a solution of 50 mM phosphate buffer (pH 11.5), 50 mM NaCl and 0.1% glycerol. The obtained solution was adjusted to pH 7.0 using 3.0 M HCl and then concentrated. The sample was dialyzed against 30 mM phosphate buffer, pH 7.0, containing 50 mM NaCl and loaded onto a Superdex 200 preparative grade column equilibrated with the same buffer. The eluate was separated in two peaks that corresponded to proteins of 13 kDa and 41 kDa molecular masses [11]. The 13 kDa fraction was concentrated and dialyzed against 25 mM phosphate buffer (pH 7.0) containing 50 mM NaCl, applied to a Superdex 200 10/300 GL column equilibrated with the same buffer and concentrated by ultra filtration. The resulting eluate was a mixture of the two forms of chicken cystatin, one of which was phosphorylated at Ser80, as mentioned above. The cystatin mixture was tested for inhibitory activity using the BapNA (N-benzoyl-DL-arginine-p-nitroanilide hydrochloride) assay [10]. The homogeneity of the samples was verified using SDS-PAGE

(results not shown). All studies were performed with preparations containing both forms of cystatin.

In this work, we used partially purified chymopapain from Sigma (St. Louis MO, USA) as a starting material for the isolation of one form of chymopapain. The main form of this protein was isolated by cation-exchange liquid chromatography, as previously described [3]. Chymopapain is effective in the digestion of proteins in the pH range of 5 to 9 [13]; to avoid autolysis, partially purified chymopapain was inactivated at 25°C using iodoacetic acid [5].

### Spectroscopic Methods

The CEW-cystatin and chymopapain concentrations were determined using the extinction coefficients  $\epsilon_{1\text{cm},280}^{1\%} = 8.7$  and 18.3, respectively [5, 14]. The ultraviolet absorption spectra of the proteins were corrected for light scattering by fitting a linear regression to a plot of  $\log(\text{absorbance})$  vs.  $\log(\text{wavelength})$  from 320 to 350 nm [15].

CD spectra were obtained in a JASCO J-715 spectropolarimeter at 25 °C in the following 50 mM buffers: formate (pH 3.5), acetate (pH 4.5), MES (pH 5.5), phosphate (pH 6.5, 7.5), Tris (pH 8.0, 8.5), and glycine (pH 9.0, 9.5, 10.0). The spectra were obtained from 100  $\mu\text{g ml}^{-1}$  protein solutions in a 0.1-cm cell. The far-UV spectra were recorded in the 190–250 nm range. The analyzed spectra were the average of three scans and were corrected by subtracting the buffer signal.

The fluorescence measurements were performed at 25°C using an ISS K2 spectrofluorometer (Urbana, USA) equipped with a water-jacketed cell holder for temperature control. Cells with 1.0-cm path lengths were used, and the samples were maintained in constant magnetic stirring. The emission spectra were recorded from 300 to 400 nm with an emission bandwidth of 8 nm, using an excitation wavelength of 280 nm and an excitation bandwidth of 16 nm. The concentrations of chymopapain and CEW-cystatin were 0.10  $\mu\text{M}$  and 0.12  $\mu\text{M}$ , respectively. Titrations of chymopapain with cystatin were monitored using measurements at the wavelength at which the maximal fluorescence change was observed (338–346 nm, depending on pH; Table 1). Two milliliters of a 0.10  $\mu\text{M}$  chymopapain solution was titrated with successive volumes (2.0–5.0  $\mu\text{L}$ ) of 9–11  $\mu\text{M}$  cystatin. Parallel titrations of 2.0 mL of buffer with the same successive volumes of cystatin were performed. Analyses were performed using the following buffers at 50 mM, with 0.1 M NaCl: MES (pH 6.0), PIPES (pH 6.5), phosphate (pH 7.0, 7.4), Tris (pH 7.2–8.8), borate (pH 9.0) and glycine (pH 9.0–10.0). In addition, we performed titrations in the following buffers, whose ionic strength was adjusted to 0.011 M using NaCl: 26 mM formate (pH 3.5), 26 mM acetate (pH 4.5), 60 mM (MES pH 5.5), 26 mM MES (pH 6.0), and 16 mM MES (pH 6.5).

### Titration Data Analysis

The binding of CEW-cystatin to chymopapain was analyzed according to the two parameter binding equation [16, 17]:

$$Y = \frac{F_P + F_C - F_M}{F_P} = \frac{a}{2} \left\{ \left( 1 + \frac{K_d}{P_T} + \frac{C_T}{P_T} \right) - \sqrt{\left( 1 + \frac{K_d}{P_T} + \frac{C_T}{P_T} \right)^2 - 4 \frac{C_T}{P_T}} \right\} \quad \text{Eq. 1}$$

Where  $F_P$  is the fluorescence of free chymopapain (considering dilution effects);  $F_C$  is the fluorescence of free chicken cystatin (parallel titration);  $F_M$  is the fluorescence of chymopapain/CEW-cystatin mixtures;  $a$  is one of the fitting parameters (the asymptotic value to which  $Y$  approaches at high values of the independent variable); the second fitting parameter is  $K_d$ , the dissociation equilibrium constant;  $P_T$  is the total chymopapain concentration;  $C_T$  is the total CEW-cystatin concentration; and the independent variable is  $(C_T/P_T)$ . The software Origin Pro 7.0 (Origin Lab Corporation) was used to perform the curve fitting.

### Modified Proton-Linkage Model

The proton-linkage model [18] can explain the change in  $K_b$  with pH by relating the binding constant with the  $pK_a$  variations of all the affected ionizing residues and the concentration of  $H^+$  ions. More recently, a modification to this model has been proposed [19] that simplifies the analysis, drastically reducing the number of parameters to be determined. This new proposal groups ionizing side chains that change their protonation state in the same pH range, the so-called *isoacidic* residues; thus, the fitted parameters are interpreted as composite  $pK_a$  constants assigned to the entire groups and not to single ionizing groups. From the equations of the method published elsewhere [19], the following equation can be obtained:

$$K_b = K_0 \frac{\sum_{i=0}^M \sum_{j=1}^{p_i} [10^{-pH}]^{-a_i^j} 10^{-\sum_{k=1}^M x_{ik}^j m_k p K_{a_k}^c}}{\sum_{i=0}^M \sum_{j=1}^{p_i} [10^{-pH}]^{-a_i^j} 10^{-\sum_{k=1}^M x_{ik}^j m_k p K_{a_k}^s}} \quad \text{Eq. 2}$$

This equation relates the binding constant ( $K_b$ ) at a certain pH to the binding constant of the fully protonated states ( $K_0$ ), the number of isoacidic groups ( $M$ ) as indexed by  $k$ , and the corresponding  $pK_a$  in the separate proteins ( $pK_a^s$ ) and in the complex ( $pK_a^c$ ). The index  $i$  represents the number of deprotonated isoacidic residues ( $i \leq M$ ),  $j$  is a protonation state index,  $a_i^j$  is the number of protons dissociated from the protein,  $x_{ik}^j$  is either 0 or 1 depending on whether site  $i$  is protonated or not, and  $m$  is the number of single residues that constitute the composite isoacidic residue (for more details, see Crnogorac *et al.*, 2001 [19]).

### Modeling of CEW-cystatin–chymopapain Dimer and Molecular Dynamics Simulations

The CEW-cystatin–chymopapain dimer was initially modeled by docking the first chicken cystatin NMR structure (PDB code 1A67) to the crystallographic structure of chymopapain (PDB code 1YAL) using the pyDockWeb server [20]; the thiomethyl moieties on Cys25 and Cys117 of chymopapain were previously removed. We used the restricted version of pyDock(RST), which imposes constraints on the possible interface contacts. The atomic model of the homologous papain-stefinB complex (PDB code 1STF) [21] and

the amino acid sequence alignments between chymopapain and papain, and cystatin and stefin indicated that the possible interface residues would be Asn18, Gln19, Gly20, Ala21, Cys22, Gly23, Cys63, Lys64, Gly65, Gly66, Tyr67 and Gln142 for chymopapain and Gly9, Ala10, Pro11, Gln53, Leu54, Val55, Ser56 and Gly57 for cystatin. The highest scoring docked models were superimposed with the crystallographic structure of the papain-stefin B complex, and their positional root mean square deviations (RMSDs) were calculated. The model with the lowest RMSD was selected as the starting structure for further refinement by molecular dynamics simulations.

Molecular dynamics (MD) simulations were performed using GROMACS 4.5.4 software [22] with the OPLS force field [23]. The side-chain ionization states in the separate proteins and in the complex at the various pH values were established using the  $pK_a$  values estimated using PROPKA [24]. The CEW-cystatin–chymopapain model was placed in the center of a periodic dodecahedral box, with 10 Å between the complex and the edge of the box. A total of 24,899 TIP3P water molecules were needed to fill the box, and 12 chloride ions were needed to neutralize the net protein charge. Prior to MD simulations, the system was relaxed by energy minimization, followed by 100 ps of thermal equilibration under the position restraints of protein heavy atoms through a harmonic force constant of 1000 kJ mol<sup>-1</sup> nm<sup>-1</sup>. MD simulation was performed using an NPT ensemble at 300 K and 1.0 bar for 100 ns. A LINCS algorithm was applied to constrain the length of all covalent bonds [25], and a 2 fs time step was used. A cutoff of 1.2 nm was applied for short-range electrostatic and van der Waals interactions, while the long-range electrostatic forces were treated using the particle mesh Ewald method [26]. The program NACCESS was used with a probe radius of 1.4 Å to calculate the percentage of solvent accessible surface for each residue in the separate proteins (%SAS<sup>s</sup>) and in the complex (%SAS<sup>c</sup>) [27].

### Calculation of the Electrostatic Binding Energy

The electrostatic binding energy ( $\Delta G_{b,elec}$ ) takes into account two components: the solvation energy ( $\Delta G_{solv}$ ) and the Coulombic energy ( $\Delta G_{coul}$ ) [24], as indicated in Eq. 3.

$$\Delta G_{b,elec} = \Delta G_{solv} + \Delta G_{coul} \quad \text{Eq. 3}$$

The solvation and Coulombic energies for the binding of chymopapain to CEW-cystatin were determined using the Adaptive Poisson-Boltzmann Solver (APBS) program [28], as described previously [29]. PROPKA [24] was used to assign the protonation state of the ionizing residues. Dielectric constants of 78.0 and 4.0 were used for water and protein, respectively. Protein parameters were obtained from the PDB2PQR server [30]. The ionic radii and atomic charges were assigned from the force field PARSE [31].

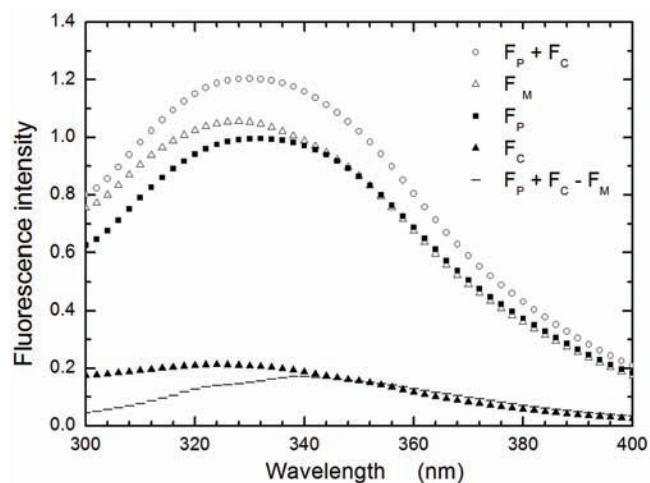
## RESULTS

### Changes in Fluorescence Accompanying Binding

The fluorescence emission spectra of CEW-cystatin, chymopapain, and a mixture of cystatin and chymopapain in



a molar ratio of 1.2 (inhibitor/protease) at pH 7.4 are shown in Figure 1. A comparison of the spectrum of the mixture with the sum of the individual spectra of the inhibitor and the protease (that would correspond to a non-interacting system) revealed that the complexation reaction was accompanied by fluorescence quenching (see Fig. 1). This result implies that the chymopapain-cystatin interaction modifies the environment of some tryptophans and produces a decrease in their quantum yield. As shown in Figure 1, the fluorescence quenching at pH 7.4, which was calculated as  $F_P + F_C - F_M$ , exhibited a maximum at 342 nm. As demonstrated previously [32], this change in fluorescence between the free and associated species allowed us to determine the binding constant ( $K_b$ ) at different pH values.

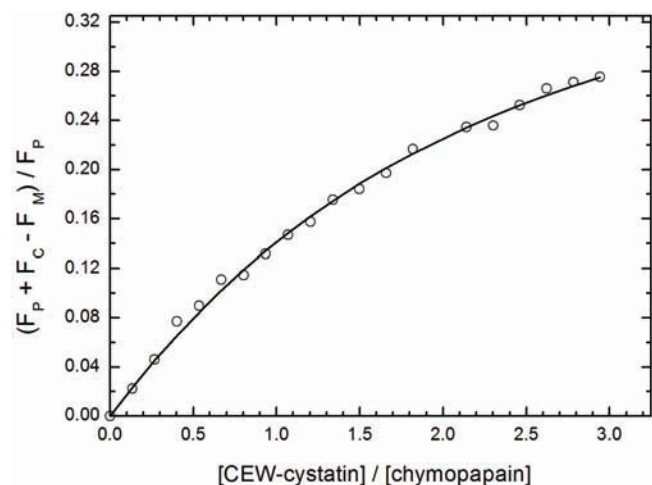


**Figure 1.** Fluorescence emission spectra of chymopapain ( $F_P$ ), CEW-cystatin ( $F_C$ ), and the chymopapain-CEW-cystatin mixture ( $F_M$ ), along with the sum of the chymopapain and CEW-cystatin spectrum ( $F_P + F_C$ ) and the difference between the free proteins and mixture spectrum ( $F_P + F_C - F_M$ ). The spectra were recorded at 25°C and at pH 7.4.

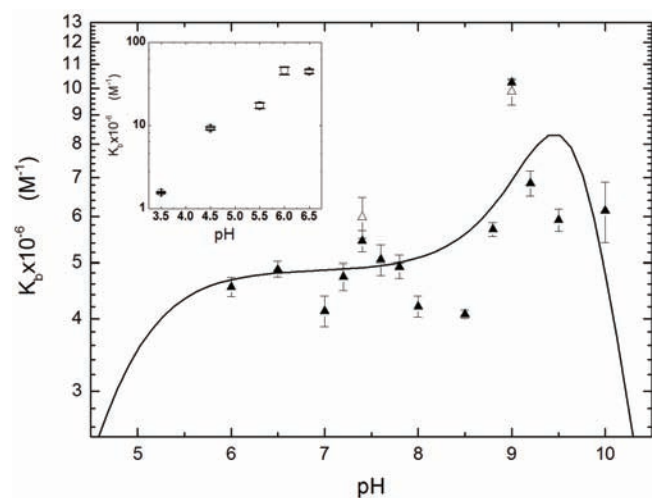
### Fluorimetric Measurements of $K_b$ as a Function of pH

The binding constant between CEW-cystatin and chymopapain,  $K_b$ , was determined using fluorimetric titrations of the protease with the inhibitor at 25°C and different pH values. A representative binding isotherm at pH 7.4 is shown in Figure 2. The line in the figure represents the fitting of Eq. 1 to the data, and one of the parameters obtained using this analysis is the dissociation constant,  $K_d$ . We then calculated  $K_b$  from  $K_b = 1/K_d$ . The results obtained at the various assayed pH values are summarized in Table 1.

The variation in the affinity of the chymopapain-cystatin dimer with pH is shown in Figure 3. This profile exhibits a maximum at pH 9.0 and three inflection points: one in the acid region that can be observed only at low ionic strength (inset in Figure 3), a second inflection point that is located in a moderately basic region before the maximum, and a third one that is apparent in the more basic region after the maximum. Using the modified proton-linkage model [19] with



**Figure 2.** Fluorimetric titration of chymopapain (0.10  $\mu$ M) with CEW-cystatin (10.7  $\mu$ M) in 50 mM Tris (pH 7.4) and 0.1 M NaCl at 25°C. The circles represent the emission measurements at 342 nm, and the line indicates the curve fit to Eq. 1.



**Figure 3.** Effect of pH on the binding equilibrium constant ( $K_b$ ) of chicken cystatin with chymopapain. The bars for the data points represent the standard error of the mean (SEM) of the experimental measurements. The continuous line represents the best fit to the modified proton-linkage model (Eq. 2). The inset shows the variation of  $K_b$  for the system at acidic pH values but in the absence of the 0.10 M NaCl that is present in the solutions of the main figure.

three isoacidic groups (corresponding to the number of observed inflection points), we could obtain a profile that was consistent with the experimental data, as shown in Figure 3 (continuous line). To obtain this curve, we first screened all possible combinations of the three isoacidic groups from Table 2, excluding group B because the  $pK_a$  values of this group did not change significantly upon binding. The only combination that yielded a profile resembling the experimental one was the profile with the isoacidic groups A, C

**Table 1.** Equilibrium constants for the binding ( $K_b$ ) of chicken cystatin to chymopapain at 25 °C.

pH	replicas	$K_b \times 10^{-6} (M^{-1})$	$r^2$	$\lambda_{max}$	Buffer
3.5	2	1.56±0.03	0.991	338	Formate*
4.5	3	9.26±0.29	0.996	338	Acetate*
5.5	2	17.36±1.36	0.995	340	MES*
6.0	3	45.97±4.69	0.994	340	MES*
6.5	5	45.08±2.62	0.960	342	MES*
6.0	3	4.54±0.17	0.995	340	MES**
6.5	2	4.87±0.15	0.992	342	PIPES**
7.0	2	4.13±0.25	0.992	344	Phosphate**
7.2	2	4.73±0.25	0.996	340	Tris**
7.4	2	5.98±0.49	0.996	342	Phosphate**
7.4	2	5.44±0.23	0.993	342	Tris**
7.6	2	5.06±0.31	0.996	332	Tris**
7.8	4	4.92±0.23	0.976	342	Tris**
8.0	2	4.20±0.18	0.994	342	Tris**
8.5	3	4.07±0.07	0.996	344	Tris**
8.8	2	5.70±0.16	0.991	342	Tris**
9.0	2	10.24±0.09	0.995	346	Glycine**
9.0	2	9.87±0.53	0.990	342	Borate**
9.2	2	6.85±0.34	0.987	340	Glycine**
9.5	2	5.92±0.26	0.989	342	Glycine**
10	2	6.14±0.73	0.990	346	Glycine**

The dissociation constant ( $K_d=K_b^{-1}$ ) and the coefficient of determination ( $r^2$ ) were obtained using non-linear regressions analysis of the titration curves, as described in Methods. The uncertainties represent the standard error of the mean (SEM) of the experimental measurements.  $\lambda_{max}$  represents the wavelength of maximum emission in the calculated difference spectrum between the free and associated species. \*Buffers at 0.01M ionic strength. \*\*Buffers at 50 mM concentration with 0.10 M NaCl.

**Table 2.** Interfacial ionizing residues in the CEW-cystatin–chymopapain dimer.

Isoacidic residue	Residue	$pK_a^s/pK_a^c$	%SAS <sup>s</sup> /%SAS <sup>c</sup>	$\Delta G_{b,elec}$ (kJ/mol)	$R^\dagger$
A	TYR61A	11.62/9.58	49.0/40.5	-255.0	0.55
B	LYS64A	10.41/10.33	86.5/61.1	-429.5	0.92
	LYS145A	10.41/10.23	69.9/51.3	-398.9	0.86
C	TYR67A	9.48/11.61	48.1/39.8	-255.0	0.55
	LYS139A	10.46/11.52	76.4/59.1	-339.3	0.73
D	GLU135A	3.58/4.3	6.3/6.3	-46.1	0.10
	ASP158A	3.04/3.5	60.7/39.6	-46.2	0.10
E	LYS156A	11.61/11.28	65.9/59.3	-396.6	0.85
F	HIS159A	6.5/4.2	9.7/2.4	51.8	-0.11
G	GLY9B	7.92/6.75	94.6/37.2	-238.4	0.51
H	GLU19B	4.6/3.42	73.3/54.5	-104.5	0.22

(Table 2) Contd....

Isoacidic residue	Residue	$pK_a^s/pK_a^c$	%SAS <sup>s</sup> /%SAS <sup>c</sup>	$\Delta G_{b,elec}$ (kJ/mol)	R <sup>†</sup>
I	ARG52B	13.1/12.93	29.8/22.9	-364.3	0.78
J	LYS59B	10.8/8.8	23.6/12.7	-445.5	0.96
K	TYR100B	9.84/11.55	20.1/10.8	-276.4	0.59

Chains A and B are chymopapain and CEW-cystatin, respectively. The superscripts refer to the separate proteins (s) and the complex (c).  $pK_a$  values were predicted using PROPKA [24]. The percentages of solvent accessible surface (%SAS) are also shown for both states and were obtained using NACCESS [27]. Binding electrostatic energies ( $\Delta G_{b,elec}$ ) for variants with each residue in the alternative protonated state. Calculations were performed for an implicit solvent model at pH 7.4 and 298 K using APBS [28].

<sup>†</sup>Ratio of  $\Delta G_{b,elec}$  to the reference value obtained for the native dimer at pH 7.4 (-466.0 kJ/mol).

and H, which possess 1, 2 and 1 individual residues, respectively. The curve shown in Figure 3 represents the best-fitted curve obtained from a non-linear regression analysis of the experimental data using Eq. 2. The maximum was not very well reproduced but reached 83% of the mean measured value and was shifted by 0.46 pH units to the basic side. The parameters achieved with this fit, *i.e.*, the  $pK_a$  of these isoacidic residues in the separate proteins and in the complex, are shown in Table 3. For the first isoacidic group ( $m_1$ ),  $pK_a^s=11.9$  and  $pK_a^c=9.30$ , which are very similar to the values associated with Tyr61A of 11.62 and 9.58, respectively. For the second isoacidic group ( $m_2=2$ ), the composite  $pK_a$  value increased from the value of the separate proteins to that of the complex ( $pK_a^s/pK_a^c=9.65/11.9$ ); these values are close to those for residue C formed by Tyr67A (9.48/11.61) and Lys139A (10.46/11.52). Finally, the last isoacidic group ( $m_3$ ) has values ( $pK_a^s/pK_a^c=4.61/3.39$ ) that agree very well with those for Glu19B (4.6/3.42).

**Table 3.** Parameters from the fitting of the modified proton-linkage model<sup>a</sup> to the experimental  $K_b$  data for 0.10 M NaCl from Table 1. The  $m_i$  values remained constant during the analysis.

Parameter	
$K_{int} \times 10^{-6}$ (M <sup>-1</sup> )	0.2923
$m_1$	1
$m_2$	2
$m_3$	1
$pK_a^{1s}$	11.90
$pK_a^{1c}$	9.30
$pK_a^{2s}$	9.65
$pK_a^{2c}$	11.90
$pK_a^{3s}$	4.61
$pK_a^{3c}$	3.39

<sup>a</sup> Ref. [19]

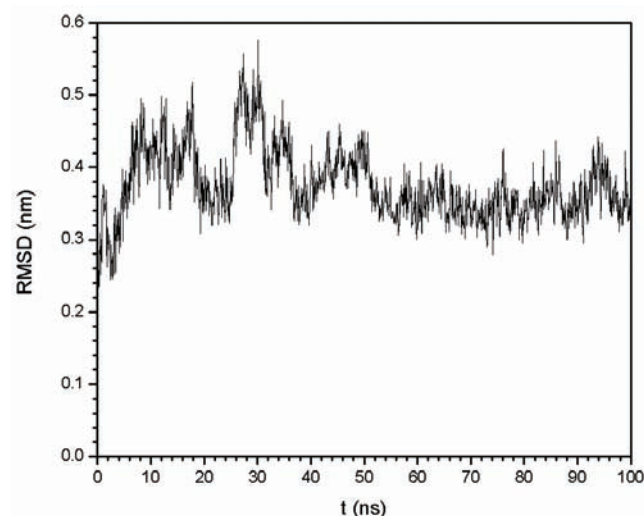
### Circular Dichroism Spectra

The far-UVCD spectra for both proteins, chymopapain and chicken cystatin, did not show significant changes at the various pH values that were assayed for binding (results not

shown). Thus, we ruled out the existence of substantial conformational changes for both proteins in the pH range between 4.0 and 10.0.

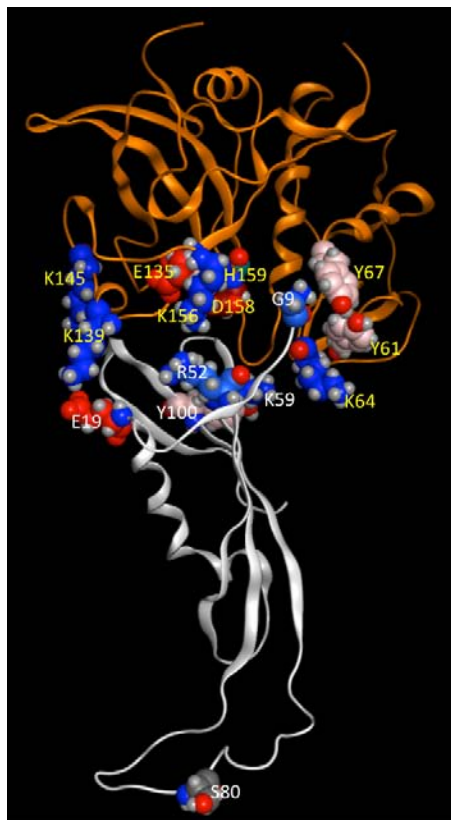
### Modeled CEW-cystatin–chymopapain Dimer

MD simulation in explicit water has been proven to be a successful strategy to improve the quality of protein molecular modeling [33]. The simulation was initiated with the enzyme-inhibitor dimer modeled by docking, and this structure was used as a reference; the resulting RMSD only reached a plateau after 52 ns (Fig. 4). The final structure of the enzyme-inhibitor complex after 100 ns of MD simulation is shown in Figure 5. We identified all the ionizing residues that were located within 5.0 Å from any atom of the neighboring chain. Table 2 presents the  $pK_a$  for all 14 of these interfacial residues in the separate proteins ( $pK_a^s$ ) and in the complex ( $pK_a^c$ ), as predicted by PROPKA [24]; the residues are grouped according to similarity in  $pK_a$  values and how this parameter changes (ascending or descending) when the proteins bind each other. The percentage of solvent accessible surface for each residue in the separated proteins (%SAS<sup>s</sup>) and in the complex (%SAS<sup>c</sup>) is also shown. A striking feature of this complex is the interchain ionic pair



**Figure 4.** Positional root mean square deviations (RMSDs) of the backbone for the CEW-cystatin–chymopapain dimer during a MD simulation in explicit solvent in an NPT ensemble at 300 K and 1 bar. The reference structure for the calculation was the initial structure.

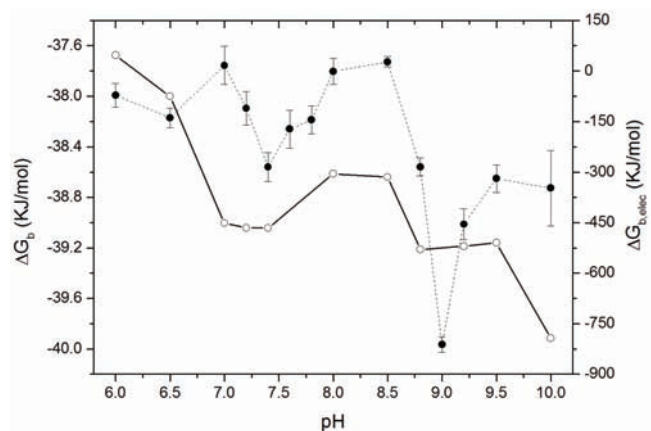
formed between Lys139 of chymopapain and Glu19 of CEW-cystatin; the distance between the NZ atoms of Lys139A and OE2 of Glu19B is 2.83 Å in the model. It should be noted also that Ser80 from cystatin, which is phosphorylated in some inhibitor molecules, is located far from the binding site of the protease on the opposite side of the molecule (Fig. 5).



**Figure 5.** Structural model of the CEW-cystatin–chymopapain dimer. This structure was obtained after a 100 ns MD simulation of the initially docked complex in explicit solvent at 300 K and 1 bar. The interfacial ionizing residues as well as the phosphorylation site of cystatin (Ser80) are highlighted with van der Waals representations. The CEW-cystatin chain and the labels of its residues are in white, while chymopapain is in orange with its residue labels in yellow.

### Electrostatic Calculations

The electrostatic binding energy ( $\Delta G_{b,elec}$ ) of the CEW-cystatin–chymopapain dimer was calculated at the different pH values at which we fluorimetrically determined the binding constant,  $K_b$ , and then the binding free energy  $\Delta G_b$  (calculated through  $\Delta G_b = -RT \ln K_b$ ). Both profiles,  $\Delta G_b$  and  $\Delta G_{b,elec}$  versus pH, present some similarity (Fig. 6), suggesting the predominance of the contribution of the electrostatic interactions to the binding process. The electrostatic binding energy calculated at pH 7.4 and 298.15 K was  $\Delta G_{b,elec} = -466.0$  kJ/mol. With this value as a reference, the significance



**Figure 6.** Comparison of the experimental binding energies obtained from the  $K_b$  values in Table 1 ( $\Delta G_b$ , filled circles) with the calculated electrostatic component ( $\Delta G_{b,elec}$ , void circles) at the various analyzed pH values. Implicit solvent calculations were performed using the APBS program [28] at 298.15 K, 0.1 M ionic strength, and dielectric constants of 78 and 4 for the solvent and the protein, respectively.

of the charged state for each ionizing interfacial residue in affecting binding was evaluated by changing the protonation state of each residue individually and recalculating the electrostatic binding energy. The results are presented in the last two columns of Table 2; the last column corresponds to the ratio ( $R$ ) of the recalculated  $\Delta G_{b,elec}$  to the reference value.

### DISCUSSION

Variations in the binding constant ( $K_b$ ) with pH for a protein dimer may be caused by the following: *i*) conformational changes in one or both proteins and *ii*) ionizing groups in the protein(s) that uptake or release  $H^+$  ions upon protein association. The circular dichroism results allow us to rule out pH-induced conformational changes in chymopapain and cystatin. Thus, we focused on the ionizing residues whose environment could change upon binding, affecting the relative stability of the protonated/unprotonated species, *i.e.*, its  $pK_a$ . The establishment of new short charge-charge interactions, either attractive or repulsive, upon protein-protein association modifies the  $pK_a$  of the involved groups [34]. Even when an interfacial charged residue does not closely approach other charges on the partner molecule, the mere change in the microenvironmental polarity does alter its stability in respect to that of the non-charged form and then induces a change in the involved  $pK_a$ .

Using a modified proton-linkage model [19], we found that the contribution of three isoacidic residues can almost completely describe the pH dependence of  $K_b$  for the CEW-cystatin–chymopapain dimer (Figure 3 and Table 3). One residue is responsible for the increase in affinity with increased pH in the acid region (insert in Figure 3), which corresponds to the  $m_3$  residue that we identified as Glu19B. This

**Table 4.** Ionizing residues and structural properties of heterodimer interfaces from cysteine proteases–cystatins complexes.

Dimer	Cystein-protease	Cystatin	N <sub>HB</sub>	N <sub>IP</sub>	Interfacial area (Å <sup>2</sup> )
Chymopapain–CEW-Cystatin	Y <sup>61</sup> , Y <sup>67</sup> , K <sup>64</sup> , E <sup>135</sup> , K <sup>139</sup> , K <sup>145</sup> , K <sup>156</sup> , D <sup>158</sup> , H <sup>159</sup>	G <sup>9</sup> , E <sup>19</sup> , R <sup>52</sup> , K <sup>59</sup> , Y <sup>100</sup> ,	6	1	1589.3
Papain–CEW-Cystatin	Y <sup>61</sup> , Y <sup>67</sup> , K <sup>139</sup> , Y <sup>144</sup> , R <sup>145</sup> , K <sup>156</sup> , D <sup>158</sup> , H <sup>159</sup>	G <sup>9</sup> , E <sup>19</sup> , R <sup>52</sup> , K <sup>59</sup> , Y <sup>100</sup>	10	1	1559.9
Papain–Human Stefin B <sup>a</sup>	Y <sup>61</sup> , Y <sup>67</sup> , K <sup>139</sup> , R <sup>145</sup> , K <sup>156</sup> , D <sup>158</sup> , H <sup>159</sup>	E <sup>15</sup> , R <sup>68</sup> , H <sup>75</sup> , E <sup>76</sup> , Y <sup>97</sup>	8	1	1965.1
Cathepsin B– Stefin A <sup>b</sup>	D <sup>22</sup> , D <sup>69</sup> , Y <sup>75</sup> , E <sup>109</sup> , H <sup>111</sup> , R <sup>116</sup> , D <sup>124</sup> , E <sup>194</sup> , H <sup>199</sup> , D <sup>224</sup> , E <sup>245</sup>	E <sup>8</sup> , K <sup>10</sup> , Y <sup>43</sup> , Y <sup>53</sup> , Y <sup>54</sup> , K <sup>68</sup> , D <sup>79</sup> ,	13	2	2639.7
Cathepsin H– Stefin A <sup>c</sup>	H <sup>61</sup> , K <sup>147</sup> , R <sup>148</sup> , D <sup>162</sup> , K <sup>163</sup> , H <sup>166</sup>	K <sup>10</sup> , E <sup>15</sup> , Y <sup>53</sup> , Y <sup>54</sup> , K <sup>68</sup> , K <sup>77</sup>	9	0	1878.3
Cathepsin L– Stefin A <sup>d</sup>	E <sup>141</sup> , D <sup>160</sup> , D <sup>162</sup> , H <sup>163</sup> , E <sup>192</sup> ,	Y <sup>54</sup> , K <sup>68</sup> , K <sup>71</sup> , E <sup>78</sup>	5	1	1864.4
Cathepsin V– Stefin A <sup>e</sup>	K <sup>20</sup> , R <sup>72</sup> , K <sup>148</sup> , K <sup>160</sup> , D <sup>163</sup> , H <sup>164</sup> , E <sup>193</sup>	E <sup>8</sup> , Y <sup>54</sup> , K <sup>68</sup> , K <sup>71</sup>	7	0	1828.1

The structure of the Chymopapain–CEW-Cystatin dimer was modeled as describe in Materials and Methods. The structure of the Papain–CEW-Cystatin complex was obtained by homology modeling using the program MOE and our Chymopapain–CEW-Cystatin model as a template. For all other complexes we used the crystallographic structures obtained from the Protein Data Bank with ID codes:<sup>a</sup> 1STF, <sup>b</sup>3K9M, <sup>c</sup>1NB5, <sup>d</sup>3KSE, <sup>e</sup>3KFQ.

N<sub>HB</sub> indicates the number of hydrogen bonds across the interface.

N<sub>IP</sub> indicates the number of ionic pairs across the interface.

**Interfacial area** calculated as the change in the solvent-accessible surface area between the heterodimer and the sum of the free proteins using the NACCESS program [27].

higher affinity can be explained by the participation of this residue in a salt bridge with Lys139A (Fig. 5): the higher the pH, the greater the proportion of the carboxylate anion (–COO<sup>–</sup>) and the stronger the interaction. If this residue is protonated, the electrostatic component of the binding energy ( $\Delta G_{b,elec}$ ) becomes less favorable, decreasing to 22% of the energy for the unprotonated form (Table 2). The second inflection point is due to the isoacidic residue  $m_2$ ; the composite  $pK_a$  of this residue increases upon binding, and the composite contains Tyr67A and Lys139A. The participation of Lys139A in the previously mentioned salt bridge is consistent with an increase in its  $pK_a$ , stabilizing the protonated form. It is worth mentioning that the positive charge of this lysine was the one that most affected  $\Delta G_{b,elec}$ , *i.e.*, of the five dimeric variants with one interfacial lysine in the unprotonated state, the calculated energy for Lys139A was the least negative of all of them, decreasing by 27% from the energy of the protonated state (Table 2). However, Tyr67A became more hidden from the solvent in the complex (Table 2), and this more hydrophobic microenvironment destabilizes the phenolate ion, favoring the protonated state with a concomitant increase in the corresponding  $pK_a$ . The remaining isoacidic group,  $m_1$ , corresponds to Tyr61A, whose  $pK_a$  decreased upon binding; this increased stability of the anionic form of Tyr can be explained by the anionic O<sup>–</sup> being a better acceptor of the hydrogen bond with the neighboring –OH of Tyr67A.

The electrostatic binding energy ( $\Delta G_{b,elec}$ ) is just one of the contributors to the total binding energy ( $\Delta G_b$ ); additional factors include the non-polar solvation energy, the loss of translational-rotational entropy and other entropic terms that are not considered in this single-conformation approximation [35]. However, the electrostatic term is important in this binding reaction, as shown by the symptomatic tenfold decrease in  $K_b$  upon increase in the environmental ionic strength from 0.011 M to 0.12 M at pH 6.0 (Table 1). The

similarity in the profiles of the pH dependence for  $\Delta G_b$  and  $\Delta G_{b,elec}$  (Fig. 6) suggest that the binding electrostatics determined using this methodology, which is based on structural information, provide the approximate relative behavior of  $\Delta G_b$  as a function of pH.

The pH dependence of the affinity between CEW-cystatin and s-(carboxymethyl)-chymopapain here presented is very different to that observed for CEW-cystatin binding s-(carboxymethyl)-papain, where the affinity decreases with pH in the 8.2–10.5 range [9]. On the other hand, quail cystatin is a much stronger human cathepsin B inhibitor ( $K_i = 47$  pM) than CEW-cystatin ( $K_i = 828$  pM) [36]; this stronger interaction was explained as due to the presence of two new possible interchain ionic pairs in the quail cystatin complex as compared with the CEW-cystatin one [36]. Certainly, the number, identity and distribution of ionizing residues at the particular proteinase–inhibitor interface will determine the electrostatic contribution to the binding energy as well as the pH dependence of its affinity. Table 4 shows how varied can be the interfaces of cysteine proteinase–inhibitor complexes, with different numbers of charged residues, hydrogen bonds, ionic pairs and interfacial area.

## CONCLUSIONS

During the association of chymopapain with chicken cystatin, electrostatic interactions play an important role. These interactions are modulated by the environmental pH, which dictates their protonation state according to the  $pK_a$  of the interfacial ionizing residues in the separate proteins and the heterodimer and thus also dictates the electrostatic binding energy. We propose that the residues involved in the interchain ionic pair (Lys139A and Glu19B from chymopapain and CEW-cystatin, respectively), as well as Tyr61A and Tyr67A from chymopapain, are the main residues responsible for the experimentally observed pH dependence of the CEW-cystatin–chymopapain dimer affinity.

## CONFLICT OF INTEREST

The authors confirm that this article content has no conflict of interest.

## ACKNOWLEDGEMENTS

This work was supported by the National Council of Science and Technology (CONACyT, México) by grants numbers 181049 and 168692 and by scholarship number 28595 to F.R.

R.A.Z. acknowledges the Centro de Supercomputo of Universidad Autónoma Metropolitana-Iztapalapa for providing computer time.

## REFERENCES

- [1] Li, X.; Liu, Z.; Cheng, Z.; Cheng, X. Cysteine cathepsins: multifunctional enzymes in cardiovascular disease. *Chonnam. Med. J.*, **2012**, *48*, 77-85.
- [2] Palermo, C.; Joyce, J.A. Cysteine cathepsin proteases as pharmacological targets in cancer. *Trends Pharmacol. Sci.*, **2007**, *29*, 22-28.
- [3] Solís-Mendiola, S.; Arroyo-Reyna, A.; Hernández-Arana, A.; Circular dichroism of cysteine proteinases from papaya latex. Evidence of differences in the folding of their polypeptide chains. *Biochim. Biophys. Acta*, **1992**, *1118*, 288-292.
- [4] Brocklehurst, K.; Baines, B.S.; Kierstan, M.P.J. Papain and other constituents of *Carica papaya* L. *Topics in Enzyme and Fermentation Biotechnology*, **1981**, *5*, 262-335.
- [5] Anastasi, A.; Brown, M.A.; Kembhav, A.A.; Nicklin, M.J.; Sayers, C.A.; Sunter, D.C.; Barrett, A.J. Cystatin, a protein inhibitor of cysteine proteinases. Improved purification from egg white, characterization, and detection in chicken serum. *Biochem. J.*, **1983**, *211*, 129-138.
- [6] Bode, W.; Engh, R.; Musil, D.; Thiele, U.; Huber, R.; Karshikov, A.; Brzin, J.; Kos, J.; Turk, V. The 2.0 Å X-ray crystal structure of chicken egg white cystatin and its possible mode of interaction with cysteine proteinases. *EMBO J.*, **1988**, *7*, 2593-2599.
- [7] Laber, B.; Krieglstein, K.; Henschen, A.; Kos, J.; Turk, V.; Huber, R.; Bode, W. The cysteine proteinase inhibitor chicken cystatin is a phosphoprotein. *FEBS Lett.*, **1989**, *248*, 162-168.
- [8] Lindhal, P.; Alriksson, E.; Jörnval, H.; Björk, I. Interaction of the cysteine proteinase inhibitor chicken cystatin with papain. *Biochemistry*, **1988**, *27*, 5074-5082.
- [9] Björk, I.; Ylinenjärvi, K. Interaction of chicken cystatin with inactivated papains. *Biochem. J.*, **1989**, *260*, 61-68.
- [10] Barrett, A.J. Cystatin, the egg white inhibitor of cysteine proteinases. *Methods Enzymol.*, **1981**, *80*, 771-778.
- [11] Trziszka, T.; Saleh, Y.; Kopece, W.; Siewiński, M.; Wesierska, E. Effect of hen's age on the level of cystatin in the chicken egg white. *Int. J. Poult. Sci.*, **2004**, *3*, 471-477.
- [12] Saleh, Y.; Wnukiewicz, J.; Andrzejak, R.; Trziszka, T.; Siewiński, M.; Ziolkowski, P.; Kopece, W. Cathepsin B and Cysteine Protease Inhibitors in Human Tongue Cancer. *J. Cancer Mol.*, **2006**, *2*, 67-72.
- [13] Henderson, P. A. Linear equation that describes the steady-state kinetics of enzymes and subcellular particles interacting with tightly bound inhibitors. *Biochem. J.*, **1972**, *127*, 312-333.
- [14] Zucker, S.; Buttle, D.J.; Nicklin, M.J.; Barrett, A.J. The proteolytic activities of chymopapain, papain, and papaya proteinase III. *Biochim. Biophys. Acta*, **1985**, *828*, 196-204.
- [15] Levine, R.L.; Federici, M.M. Quantitation of aromatic residues in proteins: Model compounds for second-derivative spectroscopy. *Biochemistry*, **1982**, *21*, 2600-2606.
- [16] Krause, J.; Bühner, M.; Sund, H. Studies of glutamate dehydrogenase. The binding of NADH and NADHP to beef-liver glutamate dehydrogenase. *Eur. J. Biochem.*, **1974**, *41*, 593-602.
- [17] Arroyo-Reyna, J.A.; Gutiérrez-Gómez, E. Affinity of chicken cystatin for S-(carbamoylmethyl)-papain, measured by fluorescence at acidic pH. *Protein Pept. Lett.*, **2002**, *9*, 231-235.
- [18] Wyman, J. Linked Functions and Reciprocal Effects in Hemoglobin: A Second Look. *Adv. Prot. Chem.*, **1964**, *19*, 223-286.
- [19] Crnogorac, M.M.; Ullmann, G.M.; Kostić, N.M. Effects of pH on Protein Association: Modification of the Proton-Linkage Model and Experimental Verification of the Modified Model in the Case of Cytochrome c And Plastocyanin. *J. Am. Chem. Soc.*, **2001**, *123*, 10790-10798.
- [20] Jimenez-García, B.; Pons, C.; Fernández-Recio, J. pyDockWEB: a web server for rigid-body protein-protein docking using electrostatics and desolvation scoring. *Bioinformatics*, **2013**, *29*, 1698-1699.
- [21] Stubbs, M.T.; Laber, B.; Bode, W.; Huber, R.; Jerala, R.; Lenarcic, B.; Turk, V. The refined 2.4 Å X-ray crystal structure of recombinant human stefin B in complex with the cysteine proteinase papain: a novel type of proteinase inhibitor interaction. *EMBO J.*, **1990**, *9*, 1939-1947.
- [22] Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.*, **2008**, *4*, 435-447.
- [23] Jorgensen, W.L.; Maxwell, D.S.; Tirado-Rives, J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.*, **1996**, *118*, 11225-11236.
- [24] Li, H.; Robertson, A.D.; Jensen, J.H. Very fast empirical prediction and interpretation of protein pKa values. *Proteins*, **2005**, *61*, 704-721.
- [25] Hess, B.; Bekker, H.; Berendsen, H.J.C.; Fraaije, J.G.E.M. LINCS: a linear constraint solver for molecular simulations. *J. Comput. Chem.*, **1997**, *18*, 1463-1472.
- [26] Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An N\*log(N) method for Ewald sums in large systems. *J. Chem. Phys.*, **1993**, *98*, 10089-10092.
- [27] Hubbard, S.J.; Thornton, J.M. 'NACCESS', Computer Program, Department of Biochemistry and Molecular Biology, University College London. **1993**. <http://www.bioinf.manchester.ac.uk/naccess/>
- [28] Baker, N.A.; Sept, D.; Joseph, S.; Holst, M.J.; McCammon, J.A. Electrostatics of nanosystems: Application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. USA*, **2001**, *98*, 10037-10041.
- [29] Lira-De León, K. I.; García-Gutiérrez, P.; Serratos, I. N.; Palomera-Cárdenas, M.; Figueroa-Corona, M. P.; Campos-Peña, V.; Meraz-Ríos, M. A. Molecular mechanism of Tau aggregation induced by anionic and cationic dyes. *J. Alzheimers Dis.*, **2013**, *35*, 319-334.
- [30] Dolinsky, T.J.; Nielsen, J.E.; McCammon, J.A.; Baker, N.A. PDB2PQR: an automated pipeline for the setup, execution, and analysis of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res.*, **2004**, *32*, W665-W667.
- [31] Sitkoff, D.; Sharp, K.A.; Honig, B. Accurate calculation of hydration free energies using macroscopic solvent models. *J. Phys. Chem.*, **1994**, *98*, 1978-1988.
- [32] Björk, I.; Ylinenjärvi, K. Interaction between chicken cystatin and the cysteine proteinases actinidin, chymopapain A, and ficin. *Biochemistry*, **1990**, *29*, 1770-1776.
- [33] Fan, H.; Mark, A.E. Refinement of homology-based protein structure by molecular dynamics simulation techniques. *Prot. Sci.*, **2004**, *13*, 211-220.
- [34] Pace, C.N.; Grimsley, G.R.; Martin Scholtz J. M. Protein ionizable groups: pK values and their contribution to protein stability and solubility. *J. Biol. Chem.*, **2009**, *284*, 13285-13289.
- [35] Wong, C.F.; Hünenberger, P.H.; Akamine, P.; Narayana, N.; Diller, T.; McCammon, J.A.; Taylor, S.; Xuong, N. H. Computational analysis of PKA-Balanol interactions. *J. Med. Chem.*, **2001**, *44*, 1530-1539.
- [36] Gerhartz, B.; Engh, R.A.; Mentele, R.; Eckerskorn, C.; Torquato, R.; Wittmann, J.; Kolb, H.J.; Machleidt, W.; Fritz, H.; Auerswald, E. A. Quail cystatin: Isolation and characterisation of a new member of the cystatin family and its hypothetical interaction with cathepsin B. *FEBS Lett.*, **1997**, *412*, 551-558.