



HAL
open science

Bases moléculaires et structurales de l'interaction entre deux calréticulines de parasite et la protéine humaine C1q

Christophe J Moreau

► **To cite this version:**

Christophe J Moreau. Bases moléculaires et structurales de l'interaction entre deux calréticulines de parasite et la protéine humaine C1q. Biochimie, Biologie Moléculaire. Université Grenoble Alpes, 2015. Français. NNT : 2015GREAV021 . tel-02295182

HAL Id: tel-02295182

<https://theses.hal.science/tel-02295182>

Submitted on 24 Sep 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ GRENOBLE ALPES

Spécialité : **BIOLOGIE STRUCTURALE ET NANOBIOLOGIE**

Arrêté ministériel : 7 août 2006

Présentée par **Christophe MOREAU**

Thèse dirigée par **Christine Gaboriaud**

Co-dirigée par **Nicole Thielens**

Préparée au sein du groupe **Réponse Immunitaire aux Pathogènes et au Soi-Altéré (IRPAS)** de l'**Institut de Biologie Structurale (IBS)**

dans l'**École Doctorale Chimie et Sciences du Vivant**

Bases moléculaires et structurales de l'interaction entre deux calréticulines de parasite et la protéine humaine C1q

Thèse soutenue publiquement le **1^{er} octobre 2015**
devant le jury composé de :

Mme Christelle Breton

Professeur, Université Grenoble Alpes/CERMAV, Présidente du jury

Mme Claudine Mayer

Professeur, Université Paris-Diderot, Paris 7/Institut Pasteur, Rapporteur

Mme Lubka Roumenina

Chargée de recherche, Université Pierre et Marie Curie/CRC Paris
Rapporteur

M Régis Daniel

Directeur de Recherche, Université d'Evry-Val d'Essonne/LAMBE,
Examineur

M David Pignol

Directeur de Recherche, Université Aix-Marseille/IBEB-CEA Cadarache,
Examineur

Mme Christine Gaboriaud

Directeur de Recherche, Université Grenoble Alpes/IBS,
Directrice de thèse



Remerciements

Je tenais à remercier toutes les personnes qui m'ont permis de réaliser cette thèse et qui m'ont soutenu durant ces trois années pleines de doutes, de remises en question, de fatigue, mais aussi d'enrichissement personnel et de bonne humeur.

Mes remerciements vont tout d'abord à Christine Gaboriaud et Nicole Thielens pour avoir accepté de me faire confiance, et m'avoir accueilli au sein du laboratoire à mon retour d'Angleterre et laissé une importante liberté et indépendance pour mener à bien ce projet. Je les remercie en particulier pour le temps qu'elles m'ont accordé et leur disponibilité sans limites pour me conseiller, me faire mûrir et me guider au cours de ces trois années, mais aussi pour toutes ces discussions sans rapport avec la science autour de la machine à café, du Biacore et durant les quelques dimanches passés au synchrotron. Et merci également pour m'avoir permis de découvrir les deux parties du continent américain.

Je remercie Claudine Mayer et Lubka Roumenina de m'avoir fait l'honneur d'accepter d'évaluer ce travail. Je remercie également Régis Daniel et Christelle Breton d'avoir accepté de participer à cette soutenance comme examinateurs, ainsi que David Pignol que je remercie aussi pour avoir accepté de faire partie de mon comité de suivi de thèse et pour tous ses conseils durant ces trois années.

Je remercie également les membres du groupe IRPAS pour la bonne humeur au sein de l'équipe, même dans les moments difficiles et pour avoir été toujours disponibles pour me donner des conseils. Merci à Isa, Véro, Pascale, Evelyne pour les longues discussions et la bonne ambiance autour de la paillasse, et pour m'avoir aidé à dompter les cellules de mammifères et les tests d'activation récalcitrants. Je n'oublierai pas de continuer à parler à haute voix. Merci à Anne et JB pour m'avoir aidé à maîtriser la FPLC qui ne m'aimait pas beaucoup et pour les conseils de purifs, et Dominique pour le SAXS et les nombreuses discussions sur des sujets variés. Je tiens également à remercier toutes les personnes du groupe avec qui je n'ai pas directement travaillé mais qui ont toujours été disponibles, Catherine, Jean-Philippe, Jean-Pierre, Philippe.

Je souhaitais également remercier Juan Fontecilla et tous les membres des groupes METALLO et GSY qui m'ont accueilli lors de ma première année de thèse comme un

membre à part entière de l'équipe et ont toujours été disponibles après mon départ pour me dépanner et pour d'agréables discussions et de bons fous rires. Je voulais remercier également Adrien Favier, Isabelle Ayala, Marc Jamin, qui ont tous été disponibles au cours de nos collaborations et m'ont permis de mener à bien ce projet de thèse.

Merci à Arturo Ferreira et tous les membres de son laboratoire de m'avoir accueilli avec bonne humeur à Santiago.

Merci également aux étudiants avec qui j'ai pu passer de super moments au bureau ou à l'extérieur : Hugo, Catarina, Cécile, Coline, Rim et ses blagues venues d'ailleurs, Widade et nos grandes discussions philosophiques du vendredi soir, ainsi que ses raccourcis dans Grenoble dont elle-seule a le secret, Géraldine et ses conseils autour de nombreux cafés, Roman et Adrien pour nos discussions sportives et les tours à vélo. Merci aussi à Matthieu, Charles, Laura pour leur bonne humeur au laboratoire et les batailles de glace.

Je voulais également remercier plus particulièrement Elo pour les longues ballades m'ayant permis de découvrir la région et pour avoir tant fait pour me faire aimer Grenoble et me remettre au ski. Je n'oublie pas tous mes amis qui ont toujours été présents et m'ont aidé à tenir au cours de ces trois ans et depuis bien plus longtemps, et qui ont toujours été présents et disponibles quelle que soit la distance de Cambridge à Grenoble : merci JC, Johan, Gab, Abbas, Caro, Hélo, Tess pour toutes ces années et j'espère que ça durera encore un bout de temps.

Parce qu'une thèse et un parcours sont aussi le fruit de rencontres qui vous donnent envie de persévérer, je tenais à remercier toutes les personnes qui m'ont donné envie de poursuivre et m'ont donné cet intérêt pour la connaissance, la recherche et les sciences. Merci à Arnaud, Athanase, François, Kira et Peter pour m'avoir donné le goût de la science et pour vos conseils qu'ils soient scientifiques ou humains.

Pour finir, je voulais remercier ma famille et plus particulièrement ma mère et Denis pour m'avoir permis de réaliser ces études à travers de nombreux sacrifices et pour leur soutien sans faille en me faisant confiance et en me laissant faire mes choix et en les soutenant pleinement. Merci d'avoir supporté et subi ma fatigue et mes doutes durant ces dernières années et de m'avoir permis d'y arriver.

SOMMAIRE

| | |
|---|-------------|
| <i>LISTE DES FIGURES</i> | I |
| <i>LISTE DES TABLEAUX</i> | IV |
| NOMENCLATURE UTILISEE POUR DESIGNER LES ACIDES AMINES | VIII |
| 1. INTRODUCTION | 1 |
| <i>I. La molécule C1q</i> | 1 |
| I.1. Immunité innée et voie classique du complément | 1 |
| I.1.a. L'immunité innée | 1 |
| I.1.b. C1q dans la voie classique du système du complément | 2 |
| I.2. Aspects structuraux de C1q, ses domaines de reconnaissance et de l'activation du complexe C1 (C1q-C1r ₂ -C1s ₂) | 6 |
| I.2.a. C1q | 6 |
| I.2.b. Structure cristallographique des GRC1q | 7 |
| I.2.c. Reconnaissance des immunoglobulines | 9 |
| I.2.d. Activation des protéases C1r _s -C1s ₂ | 10 |
| I.3. Rôle de C1q dans l'efferocytose | 13 |
| I.3.a. Rôle de C1q dans la phagocytose | 14 |
| I.3.b. Implication de C1q dans le SLE et l'élimination des cellules apoptotiques | 14 |
| I.3.c. Reconnaissance par C1q de motifs moléculaires à la surface des cellules apoptotiques | 15 |
| <i>II. La calréticuline, une protéine multifonctionnelle</i> | 17 |
| II.1. Rôle de la calréticuline dans l'homéostasie du calcium | 17 |
| II.2. Contrôle qualité du repliement des protéines dans le réticulum endoplasmique : le couple calnexine-calréticuline | 17 |
| II.2.a. Le réticulum endoplasmique : siège du repliement des protéines | 17 |
| II.2.b. Le cycle calnexine-calréticuline | 19 |
| II.2.c. Structure modulaire de la calréticuline et de la calnexine | 22 |
| II.2.d. Site de fixation du glycanes dans la calréticuline | 24 |
| II.2.e. Fixation peptidique indépendante du glycanes dans la calréticuline | 28 |
| II.3. Fonctions extracellulaires de la calréticuline | 32 |
| II.3.a. Effets anti-angiogéniques | 32 |
| II.3.b. Implication dans le phénomène de cicatrisation | 32 |
| II.3.c. Interaction avec des protéines non glycosylées : une piste pour la fonction chaperonne indépendante du glycanes ? | 33 |
| II.3.d. CRT et efferocytose | 34 |
| <i>III. Stratégies d'évasion : C1q et les calréticulines de parasite</i> | 35 |
| III.1. Inhibition et subversion du système du complément | 35 |
| III.2. Reconnaissance des calréticulines de parasite par C1q | 36 |
| <i>IV. La calréticuline de Trypanosoma cruzi</i> | 37 |
| IV.1. <i>Trypanosoma cruzi</i> , agent infectieux de la maladie de Chagas | 37 |
| IV.2. <i>Trypanosoma cruzi</i> | 38 |
| IV.2.a. Cycle de vie | 38 |
| IV.2.b. Nouvelles niches écologiques : une menace persistante ? | 39 |
| IV.3. Evolution de la réponse immunitaire : de la résistance à la maladie | 40 |
| IV.3.a. Réponse du système immunitaire | 40 |
| IV.3.b. Résistance au système immunitaire | 40 |
| IV.3.c. La réponse immunitaire, acteur majeur de la phase chronique ? | 42 |
| IV.4. Invasion par dégradation tissulaire | 42 |
| IV.4.a. Interaction à la surface de la cellule hôte | 42 |
| IV.4.b. Remodelage et destruction de la matrice extracellulaire | 43 |
| IV.5. Mécanismes d'invasion non destructifs | 44 |
| IV.5.a. Le modèle dépendant des lysosomes | 44 |
| IV.5.b. Le modèle phagocytose : implication de la calréticuline | 45 |
| IV.6. Rôle de l'interaction C1q/calréticuline dans le modèle de phagocytose | 46 |
| IV.6.a. Organisation modulaire et localisation de la calréticuline | 46 |
| IV.6.b. Mécanisme d'infectivité : rôle des anticorps | 47 |
| IV.6.c. Mécanisme d'infection : un mimétisme avec les cellules apoptotiques ? | 47 |

| | |
|--|-----------|
| <i>V. Rôle de la calréticuline dans la phagocytose: l'exemple de la calréticuline d'Entamoeba histolytica</i> | 50 |
| V.1. <i>Entamoeba histolytica</i> , parasite responsable de l'amibiase..... | 50 |
| V.2. Adhésion à la surface des cellules | 52 |
| V.3. Survie à l'interface hôte-pathogène et évolution vers la maladie | 52 |
| V.3.a. Contact avec les tissus et mise en place de la réponse immunitaire..... | 54 |
| V.3.b. Evasion du système immunitaire..... | 55 |
| V.4. Elimination des cellules..... | 56 |
| V.5. Rôle de l'interaction C1q/calréticuline dans la phagocytose par l'amibe | 58 |
| V.5.a. Identification, séquençage et localisation de la calréticuline EhCRT..... | 58 |
| V.5.b. Rôle de la EhCRT à la cupule phagocytaire | 59 |
| PROJET DE THESE | 60 |
| 2. MATERIELS ET METHODES | 62 |
| I. <i>Matériel génétique et milieux de cultures</i> | 62 |
| II. <i>Méthodes de biologie moléculaire</i> | 66 |
| II.1. Amplification d'une séquence d'intérêt par PCR..... | 66 |
| II.1.a. Principe..... | 66 |
| II.1.b. Protocole et oligonucléotides utilisés pour l'amplification de la séquence codant pour les chaînes B et C des GRC1q dans le plasmide d'expression en cellule de mammifère..... | 67 |
| II.1.c. Purification de la séquence amplifiée..... | 69 |
| II.2. Délétion/Insertion de séquence d'intérêt par PCR | 69 |
| II.2.a. Principe..... | 69 |
| II.2.b. Protocole utilisé pour l'insertion de la séquence codant pour les chaînes B et C des GRC1q dans le plasmide pcDNA3.1-GRC1q_A..... | 70 |
| II.2.c. Protocole utilisé pour les délétions introduites dans les calréticulines..... | 72 |
| II.2.d. Sélection des ADN plasmidiques néosynthétisés..... | 74 |
| II.3. Oligonucléotides | 74 |
| II.4. Extraction de l'ADN plasmidique | 76 |
| II.5. Quantification et contrôle de l'ADN..... | 77 |
| II.6. Séquençage | 77 |
| II.7. Transformation des cellules bactériennes..... | 78 |
| III. <i>Production des protéines recombinantes</i> | 78 |
| III.1 Principe de la surproduction de protéines recombinantes..... | 79 |
| III.2. Protocole de surproduction des fragments de calréticuline..... | 80 |
| III.3. Protocole de surproduction des fragments de calréticuline pour la RMN | 81 |
| III.4. Protocole de surproduction des GRC1q recombinantes | 81 |
| IV. <i>Purification de protéines</i> | 82 |
| IV.1. Principe de la purification de protéines..... | 82 |
| IV.2. Protocole de purification des fragments de calréticuline | 85 |
| IV.2.a. Extraction par lyse des cellules..... | 85 |
| IV.2.b. Chromatographie d'affinité d'ion métallique (IMAC) | 85 |
| IV.2.c. Chromatographie d'exclusion de taille | 85 |
| IV.2.d. Concentration des fragments de calréticuline | 86 |
| IV.3. Protocole de purification de C1q et des GRC1q dérivées du sérum et recombinantes..... | 86 |
| IV.3.a. Extraction plasmatique | 87 |
| IV.3.b. GRC1q exprimée s en cellules de mammifères | 89 |
| IV.4. Analyse des échantillons protéiques..... | 90 |
| IV.4.a. Quantification de la concentration en protéine | 90 |
| IV.4.b. Electrophorèse sur gel de polyacrylamide (SDS-PAGE) | 90 |
| IV.4.c. Transfert sur membrane de nitrocellulose et immunorévélation (Western Blot)..... | 92 |
| IV.4.d. Spectrométrie de masse | 93 |
| IV.4.e. Séquençage N-terminal des protéines | 94 |
| IV.4.f. Expérience de thermofluorescence | 94 |
| IV.4.g. Expérience de dichroïsme circulaire..... | 95 |
| IV.4.h. Spectroscopie de fluorescence..... | 96 |
| V. <i>Méthodes d'analyse structurale des protéines étudiées</i> | 97 |
| V.1. Détermination de structure atomique de protéine par cristallographie aux rayons X..... | 97 |
| V.1.a. La cristallogénèse | 97 |
| V.1.b. Congélation des cristaux de protéines | 101 |
| V.1.c. Enregistrement des clichés de diffraction | 102 |
| V.1.d. Traitement des clichés de diffraction..... | 105 |
| V.1.e. Résolution de la structure cristallographique | 109 |
| V.1.f. Amélioration du modèle et des cartes de densité | 113 |

| | |
|--|------------|
| V.1.g. Logiciels de visualisation | 116 |
| V.2. Diffusion des rayons X aux petits angles (SAXS)..... | 117 |
| V.2.a. Principe de la diffusion des rayons X aux petits angles | 117 |
| V.2.b. Enregistrement des données | 117 |
| V.2.c. Traitement des données : Primus..... | 119 |
| V.2.d. Génération d'enveloppe <i>in silico</i> | 122 |
| V.2.e. Structure cristallographique vs Structure en solution..... | 123 |
| V.3. Modélisation de la dynamique de la protéine..... | 123 |
| V.4. Résonance Magnétique Nucléaire..... | 124 |
| V.4.a. Principe..... | 124 |
| V.4.b. Enregistrement des spectres 1D ¹ H et 2D ¹ H ¹⁵ N..... | 124 |
| VI. Analyse biochimique des interactions entre les protéines étudiées | 125 |
| VI.1. ELISA | 125 |
| VI.1.a. Principe | 125 |
| VI.1.b. Protocole de mesure de l'interaction entre les fragments de calréticulines et les régions de C1q | 126 |
| VI.1.c. Mise en place d'un protocole testant l'inhibition de la voie classique du complément par les différents fragments de calréticulines..... | 127 |
| VI.2. Résonance plasmonique de surface (SPR)..... | 129 |
| VI.2.a. Immobilisation des ligands sur la surface..... | 131 |
| VI.2.b. Injection des analytes | 133 |
| VI.2.c. Détermination des paramètres cinétiques des interactions | 133 |
| 3. RESULTATS PRINCIPAUX | 134 |
| CHAPITRE 1 : ANALYSE STRUCTURALE DES CALRETICULINES DE <i>T. CRUZI</i> ET <i>E. HISTOLYTICA</i> | 134 |
| I. Purification des fragments de calréticulines..... | 134 |
| I.1. Stratégie de dissection..... | 134 |
| I.2. Production des fragments de TcCRT..... | 136 |
| I.3. Purification des fragments de TcCRT..... | 137 |
| I.3.a. Identification du monomère | 140 |
| I.3.b. Les domaines P et C-terminaux cruciaux pour la solubilité et le repliement de la TcCRT ? | 141 |
| I.3.c. Thermofluorescence : pas uniquement un problème de solubilité..... | 142 |
| I.3.d. Dichroïsme circulaire et fluorescence : identification des fragments stables..... | 143 |
| II. Cristaux des calréticulines de <i>T. cruzi</i> et <i>E. histolytica</i> | 146 |
| II.1. Echec de la cristallisation: la calréticuline de <i>T. cruzi</i> , une protéine toujours trop flexible ? | 146 |
| II.1.a. Analyse SAXS : vers la dissection gagnante..... | 146 |
| II.1.b. Premiers cristaux, un nouveau challenge..... | 148 |
| II.1.c. Interaction avec le glucose : amélioration des données..... | 149 |
| II.2. Cristaux de la calréticuline d' <i>E. histolytica</i> (EhCRT) : utilisation des acquis?..... | 151 |
| III. Comparaison structurale des calréticulines de parasites et de mammifères..... | 155 |
| III.1. Organisation des TcCRT et EhCRT dans le cristal | 155 |
| III.1.a. Empilement de la TcCRT | 155 |
| III.1.b. Empilement de la EhCRT..... | 156 |
| III.2. Comparaison des domaines globulaires..... | 157 |
| III.2.a. Un amas de petites différences | 158 |
| III.2.b. Site lectine et fixation du glucose | 159 |
| III.2.c. Site de fixation du calcium et du chlore | 165 |
| III.3. Etude de la dynamique du domaine P et de l'hélice C-terminale | 166 |
| III.3.a. Flexibilité de l'extrémité de l'hélice C-terminale | 166 |
| III.3.b. Flexibilité du domaine P..... | 173 |
| III.3.c. Stabilité de la jonction du domaine P..... | 175 |
| III.4. Stabilisation du domaine P dans le cristal : mise en évidence d'un site d'interaction des peptides indépendants du glycane ? | 179 |
| CHAPITRE 2 : LES CALRETICULINES COMME RECEPTEURS DE C1Q | 183 |
| IV. Caractérisation des interactions entre les fragments de CRT et C1q | 183 |
| IV.1. Identification des domaines impliqués dans l'interaction..... | 183 |
| IV.2. Paramètres cinétiques de l'interaction entre C1q et les CRTs..... | 188 |
| IV.3. Importance du site lectine | 195 |
| IV.4. Vers une identification des acides aminés clés impliqués dans l'interaction..... | 197 |
| IV.5. Vers l'analyse structurale de l'interaction entre la TcCRT et les GRC1q par RMN..... | 199 |
| IV.5.a. Caractérisation de la TcCRT-N ¹⁵ | 199 |
| IV.5.b. Etudes préliminaires du complexe TcCRT-GRC1q..... | 201 |
| CHAPITRE 3 : DEVELOPPEMENT DE LA PREMIERE FORME RECOMBINANTE DES REGIONS GLOBULAIRES DE C1Q | 205 |
| V. Construction du plasmide d'expression pcDNA3.1-GRC1q..... | 205 |

| | |
|---|-------------------|
| <i>VI. Expression en cellules de mammifères.....</i> | <i>206</i> |
| <i>VII. Purification des GRC1q recombinantes.....</i> | <i>208</i> |
| <i>VIII. Caractérisation des GRC1q recombinantes</i> | <i>209</i> |
| VIII.1. Vérification de la masse et de la séquence N-terminale de la protéine | 209 |
| VIII.2. Contrôle qualité RMN 1D..... | 210 |
| VIII.3. Etudes fonctionnelles des GRC1q recombinantes | 210 |
| VIII.4. Etude structurale des GRC1q recombinantes | 211 |
| 4. DISCUSSION | 218 |
| <i>I. Forme monocaténaire des régions globulaires de C1q.....</i> | <i>218</i> |
| <i>II. Etudes des calréticulines de T. cruzi et E. histolytica.....</i> | <i>218</i> |
| I.1. Analyses structurales..... | 218 |
| II.2. Comparaison des structures en lien avec la fonction chaperonne | 222 |
| <i>III. Analyses d'interaction in vitro avec C1q.....</i> | <i>223</i> |
| 5. COLLABORATIONS ET PERSPECTIVES FONCTIONNELLES | 233 |
| <i>BIBLIOGRAPHIE</i> | <i>235</i> |
| <i>ANNEXES</i> | <i>A</i> |

LISTE DES FIGURES

INTRODUCTION

| | |
|--|----|
| FIGURE 1.I.1: REPRESENTATION SCHEMATIQUE DE LA CASCADE DU COMPLEMENT, ACTIVEE PAR LES VOIES CLASSIQUE ET ALTERNE | 5 |
| FIGURE 1.I.2: ORGANISATION STRUCTURALE DE C1q | 6 |
| FIGURE 1.I.3: STRUCTURE CRISTALLOGRAPHIQUE DES GRC1 | 8 |
| FIGURE 1.I.4: MODELISATION DE LA SURFACE D'INTERACTION ENTRE LES GRC1Q ET LES IgG | 10 |
| FIGURE 1.I.5: ORGANISATION MODULAIRE DES PROTEASES C1r ET C1s | 11 |
| FIGURE 1.I.6: DIMERISATION DES PROTEINES C1r ET C1s | 11 |
| FIGURE 1.I.7: MODELISATION DU POSITIONNEMENT DES PROTEASES C1r ET C1s DANS LE COMPLEXE C1 | 13 |
| FIGURE 1.I.8: SITE DE FIXATION DU DEOXY-D-RIBOSE DE LA PHOSPHATIDYL-SERINE, 2 SIGNAUX EAT-ME ET MECANISMES D'ACTIVATION DU COMPLEXE C1 | 16 |
| FIGURE 1.II.1: PRISE EN CHARGE DES PROTEINES EN COURS DE SYNTHESE DANS LE RIBOSOME | 20 |
| FIGURE 1.II.2: DESCRIPTION DU CYCLE CALNEXINE-CALRETICULINE | 22 |
| FIGURE 1.II.3: STRUCTURE MODULAIRE DE LA CALRETICULINE ET DE LA CALNEXINE | 24 |
| FIGURE 1.II.4: CARTOGRAPHIE STRUCTURALE DU SITE DE RECONNAISSANCE DU CARBOHYDRATE DANS LE SITE LECTINE DE LA CALRETICULINE | 26 |
| FIGURE 1.II.5: SUPERPOSITION DU SITE LECTINE DE LA CALRETICULINE HUMAINE ET DE LA CALNEXINE DE CHIEN | 27 |
| FIGURE 1.II.6: SITE DE FIXATION DU Ca^{2+} DANS LES CALRETICULINES DE MAMMIFERE | 28 |
| FIGURE 1.II.7: SITE POTENTIEL D'INTERACTION DU PEPTIDE SUR LE BORD DU DOMAINE LECTINE DE LA CALRETICULINE | 30 |
| FIGURE 1.II.8: MODELES DE FIXATION DES PROTEINES NEOSYNTHESEES AVEC LA CALRETICULINE | 31 |
| FIGURE 1.II.9: REGIONS POTENTIELLES D'INTERACTIONS PEPTIDES-DEPENDANTES | 33 |
| FIGURE 1.IV.1: CYCLE DE VIE DU PARASITE <i>TRYPANOSOMA CRUZI</i> | 38 |
| FIGURE 1.IV.2: PREMIER MODELE D'INHIBITION DE LA VOIE CLASSIQUE DU COMPLEMENT PAR LA TCCRRT | 41 |
| FIGURE 1.IV.3: MODELE DE PENETRATION DE <i>T. CRUZI</i> DEPENDANT DES LYSOSOMES | 45 |
| FIGURE 1.IV.4: MECANISME D'INFECTION DE <i>T. CRUZI</i> : EXEMPLE DES CELLULES DU PLACENTA | 48 |
| FIGURE 1.IV.5: MECANISME D'INVASION DE TYPE « CHEVAL DE TROIE » DE <i>T. CRUZI</i> DANS LES MACROPHAGES | 49 |
| FIGURE 1.V.1: CYCLE DE VIE DE <i>ENTAMOEBE HISTOLYTICA</i> | 51 |
| FIGURE 1.V.2: REPOSE IMMUNITAIRE LORS DE LA COLONISATION DES TISSUS | 53 |
| FIGURE 1.V.3: ILLUSTRATION DES MECANISMES UTILISES PAR <i>E. HISTOLYTICA</i> POUR DETRUIRE LES CELLULES | 58 |
| FIGURE 1.V.4: PARTICIPATION DE LA CRT A LA PHAGOCYTOSE AMEBIQUE | 60 |
| FIGURE 1.VI.1: ILLUSTRATION DU PROJET DE THESE | 61 |

MATERIELS ET METHODES

| | |
|--|-----|
| FIGURE 2.I.1: CARTE VECTORIELLE DES PLASMIDES D'EXPRESSION BACTERIENS UTILISES | 64 |
| FIGURE 2.I.2: CARTE VECTORIELLE DU PLASMIDE D'EXPRESSION EN CELLULES DE MAMMIFERE UTILISE | 65 |
| FIGURE 2.II.1: PRINCIPE DE LA TECHNIQUE PCR (POLYMERASE CHAIN REACTION) | 66 |
| FIGURE 2.II.2: SCHEMA DECRIVANT LA STRATEGIE D'AMPLIFICATION DU GENE CODANT POUR GRC1Q_BC AFIN D'OBTENIR UN MEGA-PRIMER | 67 |
| FIGURE 2.II.3: PRINCIPE DE L'INSERTION/DELETION D'UNE SEQUENCE PAR PCR (MUTAGENESE DIRIGEE) | 69 |
| FIGURE 2.II.4: SCHEMA PRESENTANT LE PROTOCOLE D'INSERTION DU GENE CODANT LES CHAINES A ET B GRC1Q DANS LE PLASMIDE pCDNA3.1-GRC1Q-A | 70 |
| FIGURE 2.II.5: SCHEMA PRESENTANT LA STRATEGIE UTILISEE POUR LE CHOIX DES OLIGONUCLEOTIDES POUR SUPPRIMER DES REGIONS DES DIFFERENTES CALRETICULINES PAR PCR. | 72 |
| FIGURE 2.III.1: REGULATION DE L'OPERON LACTOSE | 79 |
| FIGURE 2.III.2: PRINCIPE DE L'EXPRESSION D'UNE SEQUENCE CODANTE CLONEE DANS UN VECTEUR UTILISANT UN PROMOTEUR T7 SOUS LE CONTROLE DE L'OPERON LACTOSE | 80 |
| FIGURE 2.IV.1: PRINCIPE DES DIFFERENTES CHROMATOGRAPHIES EN PHASE LIQUIDE UTILISEES POUR LA PURIFICATION DES PROTEINES RECOMBINANTES D'INTERET | 84 |
| FIGURE 2.IV.2: SCHEMA DE MONTAGE DE LA CASSETTE POUR LE TRANSFERT SUR MEMBRANE DE NITROCELLULOSE (WESTERN BLOT). | 93 |
| FIGURE 2.V.1: REPRESENTATION SCHEMATIQUE D'UN ARRANGEMENT CRISTALLIN | 98 |
| FIGURE 2.V.2: DIAGRAMME DE PHASE D'UNE PROTEINE | 100 |
| FIGURE 2.V.3: TECHNIQUE DE LA GOUTTE SUSPENDUE | 100 |
| FIGURE 2.V.4: SCHEMAS DECRIVANT LE PHENOMENE DE DIFFRACTION SELON LA LOI DE BRAGG | 103 |

| | |
|--|-----|
| FIGURE 2.V.5: DESCRIPTION DES SYSTEMES CRISTALLINS ET RESEAUX DE BRAVAIS | 108 |
| FIGURE 2.V.6: LE FACTEUR DE STRUCTURE | 109 |
| FIGURE 2.V.7: PRINCIPE DU SAXS | 118 |
| FIGURE 2.VI.1 : PRINCIPE DE LA RESONANCE PLASMONIQUE DE SURFACE | 130 |
| FIGURE 2.VI.2 : DIFFERENTES PHASES D'UN SENSORGRAMME | 131 |

RESULTATS PRINCIPAUX

| | |
|--|-----|
| FIGURE 3.I.1 : DESCRIPTION DE LA CONSTRUCTION DE TcCRT Δ P | 135 |
| FIGURE 3.I.2 : STRATEGIE DE DISSECTION DE L'HELICE C-TERMINALE | 136 |
| FIGURE 3.I.3 : SURPRODUCTION DES FRAGMENTS DE LA TcCRT | 137 |
| FIGURE 3.I.4 : PROFIL ET ANALYSE SDS-PAGE DE L'ETAPE DE CHROMATOGRAPHIE D'AFFINITE DE LA TcCRT Δ P Δ 281 | 138 |
| FIGURE 3.I.5 : PROFIL ET ANALYSE SDS-PAGE DE L'ETAPE DE CHROMATOGRAPHIE D'EXCLUSION DE TAILLE DE LA TcCRT Δ P Δ 281 | 139 |
| FIGURE 3.I.6 : PROFIL D'ELUTION DE L'ETAPE DE CHROMATOGRAPHIE D'EXCLUSION DE TAILLE POUR LE FRAGMENT TcCRT Δ P Δ 276 | 142 |
| FIGURE 3.I.7 : MESURE DE LA STABILITE DES FRAGMENTS DE TcCRT PAR THERMOFLUORESCENCE | 143 |
| FIGURE 3.I.8 : SPECTRE DE FLUORESCENCE DES TRYPTOPHANES DES DIFFERENTS FRAGMENTS DE TcCRT | 144 |
| FIGURE 3.I.9 : SPECTRE DE DICHROÏSME CIRCULAIRE DES FRAGMENTS DE TcCRT | 145 |
| FIGURE 3.II.1 : ETUDE DE LA FLEXIBILITE DU FRAGMENT TcCRT Δ P Δ 281 PAR SAXS | 147 |
| FIGURE 3.II.2 : TRONCATION SUPPLEMENTAIRE DANS LE DOMAINE P DU FRAGMENT TcCRT Δ P Δ 288 | 148 |
| FIGURE 3.II.3 : CRISTAUX DU FRAGMENT TcCRT Δ P Δ 288 DE LA CALRETICULINE DE <i>T. CRUZI</i> | 149 |
| FIGURE 3.II.4 : PROFIL D'ELUTION DE L'ETAPE DE CHROMATOGRAPHIE D'EXCLUSION DE TAILLE SUR LA TSKG3000 DE LA EHCRT Δ P Δ 288 | 152 |
| FIGURE 3.II.5 : EXEMPLE D'UN CRISTAL OBTENU POUR LE FRAGMENT EHCRT Δ P Δ 288 | 153 |
| FIGURE 3.III.1 : ORGANISATION DES MONOMERES DE TcCRT DANS L'UNITE ASYMETRIQUE | 156 |
| FIGURE 3.III.2 : ORGANISATION DES MONOMERES DE EHCRT DANS L'UNITE ASYMETRIQUE DES DIFFERENTS GROUPES D'ESPACE | 157 |
| FIGURE 3.III.3 : ALIGNEMENT STRUCTURAL DES FRAGMENTS DE CALRETICULINES | 158 |
| FIGURE 3.III.4 : INSERTIONS ET DELETIONS MAJEURES DANS LES FRAGMENTS DE CALRETICULINES | 159 |
| FIGURE 3.III.5 : CONSERVATION DU SITE LECTINE DANS LES CALRETICULINES | 160 |
| FIGURE 3.III.6 : INTERACTION DU GLUCOSE DANS LE SITE LECTINE DE LA TcCRT | 161 |
| FIGURE 3.III.7 : IDENTIFICATION DU LIGAND PRESENT DANS LE DOMAINE LECTINE DE LA EHCRT | 162 |
| FIGURE 3.III.8 : CARTES DE DENSITE DE LA MOLECULE DE GLUCOSE DANS LE SITE LECTINE DE LA EHCRT | 163 |
| FIGURE 3.III.9 : SITE DE FIXATION DU GLUCOSE DANS LA EHCRT | 163 |
| FIGURE 3.III.10 : SITE DE FIXATION DU GLUCOSE CONSERVE ENTRE LES CALRETICULINES | 164 |
| FIGURE 3.III.11 : SITE DE FIXATION DU CALCIUM DANS LA EHCRT | 165 |
| FIGURE 3.III.12 : SITE DE FIXATION DU CHLORURE DANS LA EHCRT ET LA TcCRT | 166 |
| FIGURE 3.III.13 : FLEXIBILITE DE L'EXTREMITÉ C-TERMINALE DES CALRETICULINES | 167 |
| FIGURE 3.III.14 : STABILISATION DE L'HELICE C-TERMINALE DES CALRETICULINES PAR LE DOMAINE GLOBULAIRE | 168 |
| FIGURE 3.III.15 : COURBE DE DIFFUSION OPTIMISEE DU FRAGMENT TcCRT Δ P Δ 288 | 170 |
| FIGURE 3.III.16 : TRAITEMENT DES DONNEES SAXS SUR LES FRAGMENTS DE CALRETICULINES | 171 |
| FIGURE 3.III.17 : COMPARAISON DES COURBES DE DIFFUSION CALCULEE ET MESUREE POUR LE FRAGMENT TcCRT Δ P Δ 288 | 172 |
| FIGURE 3.III.18: COMPARAISON DES COURBES DE DIFFUSION CALCULEES A PARTIR DE MODELES ISSUS DE BUNCH ET CELLE MESUREE POUR LE FRAGMENT TcCRT Δ P Δ 288 | 173 |
| FIGURE 3.III.19 : VALIDATION D'UN MODELE DE LA TcCRT Δ P Δ 288 | 174 |
| FIGURE 3.III.20 : CONFORMATIONS DU DOMAINE P EN SOLUTION | 175 |
| FIGURE 3.III.21: ORIENTATION DE LA JONCTION ENTRE LE DOMAINE P ET LE DOMAINE GLOBULAIRE DANS LES CALRETICULINES | 176 |
| FIGURE 3.III.22: INTERACTIONS STABILISANT LA JONCTION DU DOMAINE P DANS LES CRTS | 177 |
| FIGURE 3.III.23: DESCRIPTION MOLECULAIRE DE LA FERMETURE DU DOMAINE P | 179 |
| FIGURE 3.III.24 : DESCRIPTION DE L'INTERACTION ENTRE LE DOMAINE P ET LE DOMAINE LECTINE DANS LA TcCRT | 180 |
| FIGURE 3.III.25: DESCRIPTION DES INTERACTIONS PERMETTANT LA STABILISATION DU DOMAINE P PAR LE DOMAINE LECTINE DANS LA EHCRT | 181 |
| FIGURE 3.III.26: UNE EXCEPTION : INTERACTION ENTRE LE DOMAINE GLOBULAIRE DE LA CHAÎNE C ET LE DOMAINE LECTINE DE LA CHAÎNE A DE LA EHCRT DANS L'UNITE ASYMETRIQUE | 182 |
| FIGURE 3.IV.1 : INTERACTIONS ENTRE C1Q ET LES FRAGMENTS DE CALRETICULINES (ELISA) | 184 |
| FIGURE 3.IV.2 : INTERACTIONS ENTRE LES CLR DE C1Q ET LES FRAGMENTS DE CALRETICULINES (ELISA) | 186 |

| | |
|--|-----|
| FIGURE 3.IV.3 : INTERACTIONS ENTRE LES GRC1Q ET LES FRAGMENTS DE CALRETICULINES (ELISA) | 187 |
| FIGURE 3.IV.4 : CINETIQUES D'INTERACTIONS ENTRE LES GRC1Q ET LES FRAGMENTS DE CRTS | 190 |
| FIGURE 3.IV.5 : CINETIQUES D'INTERACTIONS ENTRE C1Q ET LES FRAGMENTS DE CRTS | 193 |
| FIGURE 3.IV.6 : ESSAIS DE COMPETITIONS AVEC LE GLUCOSE | 195 |
| FIGURE 3.IV.7 : EFFETS DU GLUCOSE ET DU MANNOSE SUR L'INTERACTION C1Q-CRT | 197 |
| FIGURE 3.IV.8 : RESIDUS DE C1Q MODULANT L'INTERACTION AVEC LES CALRETICULINES | 198 |
| FIGURE 3.IV.8: CARACTERISATION DE LA TCCRTAPTERA288 PAR UNE EXPERIENCE RMN HSQC ¹⁵ N- ¹ H | 200 |
| FIGURE 3.IV.9: SPECTRE RMN HSQC N ¹⁵ -H DE LA TCCRTAPTERA288 EN COMPLEXE AVEC DIFFERENTES CONCENTRATIONS DES GRC1Q | 202 |
| FIGURE 3.IV.10: AGRANDISSEMENT DE LA REGION RANDOM-COIL DU SPECTRE RMN HSQC TROSY ¹⁵ N- ¹ H DE LA TCCRTAPTERA288 EN COMPLEXE AVEC DIFFERENTES CONCENTRATIONS DES GRC1Q PLASMATIQUES | 203 |
| FIGURE 3.IV.11: PARTICIPATION DES TRYPTOPHANES AU COMPLEXE ENTRE LA TCCRTAPTERA288 ET LES GRC1Q | 204 |
| FIGURE 3.IV.12: RESIDUS PRINCIPAUX POTENTIELLEMENT IMPLIQUES DANS LE COMPLEXE ENTRE LA TCCRTAPTERA288 ET LES GRC1Q | 204 |
| FIGURE 3.V.1: CLONAGE DE LA SEQUENCE CODANT LES GRC1Q RECOMBINANTES DANS LE PLASMIDE PCDNA3.1 | 206 |
| FIGURE 3.VI.2: TEST D'EXPRESSION DES GRC1Q RECOMBINANTES EN CELLULES HEK 293-F | 207 |
| FIGURE 3.VI.3: CONTROLE DE LA PRODUCTION DES GRC1Q RECOMBINANTES PAR LES CELLULES HEK 293-F TRANSFECTEES DE MANIERE STABLE | 207 |
| FIGURE 3.VII.1: CHROMATOGRAMME DE L'ETAPE DE PURIFICATION DES GRC1Q RECOMBINANTES | 208 |
| FIGURE 3.VII.2 : GEL BILAN DE LA PURIFICATION DES GRC1Q RECOMBINANTES | 209 |
| FIGURE 3.VIII.1: CARACTERISATION FONCTIONNELLE DES GRC1Q RECOMBINANTES. | 211 |
| FIGURE 3.VIII.2 : PHOTO DE CRISTAUX DE GRC1Q RECOMBINANTE | 212 |
| FIGURE 3.VIII.4: SUPERPOSITION DE LA STRUCTURE DES GRC1Q RECOMBINANTES ET PLASMATIQUES | 214 |
| FIGURE 3.VIII.5: DIFFERENCES OBSERVEES DANS LES STRUCTURES DE GRC1Q PLASMATIQUES ET RECOMBINANTES. | 215 |
| FIGURE 3.VIII.6: SITE DE FIXATION DU CALCIUM DANS LES GRC1Q RECOMBINANTES CRISTALLISEES EN ABSENCE DE CALCIUM | 216 |
| FIGURE 3.VIII.7: SITE DE FIXATION DU CALCIUM DANS LES GRC1Q RECOMBINANTES CRISTALLISEE EN PRESENCE DE CALCIUM | 217 |

DISCUSSION

| | |
|--|-----|
| FIGURE 4.II.1 : DIFFERENCE DANS L'ORIENTATION DE LA JONCTION ENTRE LA CNX ET LES CRTS | 220 |
| FIGURE 4.III.1 : PREMIER MODELE PROPOSE POUR L'INTERACTION ENTRE LES C1Q ET LA CRT A L'INTERFACE ENTRE LES MACROPHAGES ET LES CELLULES APOPTOTIQUES | 224 |
| FIGURE 4.III.2 : SURFACE ELECTROSTATIQUE DES CALRETICULINES (VUE DE DESSUS) | 226 |
| FIGURE 4.III.3 : INFLUENCE DE LA FLEXIBILITE DU DOMAINE P SUR LA SURFACE ELECTROSTATIQUE DE LA EHCRT | 227 |
| FIGURE 4.III.4 : SITE D'INTERACTION POTENTIEL DE C1Q DANS LA EHCRT ET LA HUCRT | 227 |
| FIGURE 4.III.5: SURFACE ELECTROSTATIQUE DES CALRETICULINES (VUE LATERALE, DOMAINE P A GAUCHE) | 229 |
| FIGURE 4.III.6 : SITE DE FIXATION DE C1Q DANS LA EHCRT ET LA TCCRT | 230 |
| FIGURE 4.III.7 : DEPLACEMENT DU TRYPTOPHANE 198/193 LORS D'UNE INTERACTION | 231 |
| FIGURE 4.III.8 : HYPOTHESE PROPOSEE POUR L'INTERACTION ENTRE C1Q ET LA CRT A L'INTERFACE ENTRE LES MACROPHAGES ET LES CELLULES APOPTOTIQUES | 232 |

ANNEXES

| | |
|--|---|
| ANNEXE 1 : ALIGNEMENT DE SEQUENCE DES CALRETICULINES ETUDIEES AU COURS DE CETTE THESE ET DE LA CALRETICULINE DE <i>B. MALAYI</i> | A |
| ANNEXE 2 : MODELE D'INTERACTION ENTRE LA CRT DE <i>B. MALAYI</i> ET LES GRC1Q OBTENU PAR MODELISATION MOLECULAIRE (YADAV ET AL., 2014) | B |
| ANNEXE 3 : SEQUENCES DES GRC1Q RECOMBINANTES PRODUITES EN CELLULES HEK 273-F ET ALIGNEMENT DE SEQUENCES ENTRE LES GRC1Q RECOMBINANTES ET LES GRC1Q PLASMATIQUES | C |
| ANNEXE 5 : ALIGNEMENT DE SEQUENCE DE CALRETICULINES | E |
| ANNEXE 6 : ALIGNEMENT DE SEQUENCE DE CALRETICULINES ET DE CALNEXINES | J |

LISTE DES TABLEAUX

MATERIELS ET METHODES

| | |
|---|-----|
| <u>TABLEAU 2.I.1</u> : TABLEAU PRESENTANT LES DIFFERENTES SOUCHES CELLULAIRES UTILISEES POUR LE CLONAGE ET L'EXPRESSION DES FRAGMENTS DE CALRETICULINE ET DES GRC1Q RECOMBINANTES | 62 |
| <u>TABLEAU 2.I.2</u> : TABLEAU PRESENTANT LES DIFFERENTS MILIEUX DE CULTURES UTILISES POUR LE CLONAGE ET L'EXPRESSION DES FRAGMENTS DE CALRETICULINE ET DES GRC1Q RECOMBINANTES | 63 |
| <u>TABLEAU 2.II.1</u> : MELANGE DE PCR POUR L'AMPLIFICATION DE LA SEQUENCE CODANT GRC1Q_BC | 68 |
| <u>TABLEAU 2.II.2</u> : PROGRAMME DE PCR POUR L'AMPLIFICATION DE LA SEQUENCE CODANT GRC1Q_BC | 68 |
| <u>TABLEAU 2.II.3</u> : MELANGE DE PCR POUR L'AMPLIFICATION DE LA SEQUENCE CODANT GRC1Q_BC | 71 |
| <u>TABLEAU 2.II.4</u> : PROGRAMME DE PCR POUR L'AMPLIFICATION DE LA SEQUENCE CODANT GRC1Q_BC | 71 |
| <u>TABLEAU 2.II.5</u> : MELANGE DE PCR POUR LES DELETIONS INTRODUITES DANS LES CALRETICULINES | 73 |
| <u>TABLEAU 2.II.6</u> : PROGRAMME DE PCR POUR LES DELETIONS INTRODUITES DANS LES CALRETICULINES | 73 |
| <u>TABLEAU 2.II.7</u> : OLIGONUCLEOTIDES UTILISES POUR L'AMPLIFICATION DE LA SEQUENCE CODANT POUR LES CHAINES B ET C DES GRC1Q (LA SEQUENCE EST DONNEE DE 5' EN 3') | 74 |
| <u>TABLEAU 2.II.8</u> : OLIGONUCLEOTIDES UTILISES POUR INTRODUIRE DES DELETIONS DANS L'HELICE C-TERMINALE DE LA CALRETICULINE DE <i>T. CRUZI</i> (LA SEQUENCE EST DONNEE DE 5' EN 3') | 75 |
| <u>TABLEAU 2.II.9</u> : OLIGONUCLEOTIDES UTILISES POUR INTRODUIRE DES DELETIONS DANS LE DOMAINE P DES CALRETICULINES DE <i>T. CRUZI</i> ET <i>E. HISTOLYTICA</i> (LA SEQUENCE EST DONNEE DE 5' EN 3') | 76 |
| <u>TABLEAU 2.V.1</u> : CONDITIONS DE CRISTALLISATION DU FRAGMENT TcCRT Δ PTER Δ 288 | 101 |
| <u>TABLEAU 2.V.2</u> : CONDITIONS DE CRISTALLISATION DU FRAGMENT EHCRT Δ PTER Δ 288 | 102 |
| <u>TABLEAUX 2.V.3</u> : SYSTEMES CRISTALLINS ET EXEMPLES DE GROUPES D'ESPACE ASSOCIES | 108 |
| <u>TABLEAU 2.V.4</u> : LISTE DES ECHANTILLONS ETUDIES PENDANT LES EXPERIENCES DE SAXS | 119 |
| <u>TABLEAU 2.VI.1</u> : CONCENTRATIONS DE CRTS ADSORBEES POUR LES TESTS D'INTERACTIONS ELISA | 127 |
| <u>TABLEAU 2.VI.2</u> : CONDITIONS D'IMMOBILISATION DES PROTEINES ETUDIEES (LA QUANTITE IMMOBILISEE DE HUCRT EST UTILISEE COMME REFERENCE) | 132 |

RESULTATS PRINCIPAUX

| | |
|---|-----|
| <u>TABLEAU 3.I.1</u> : CARACTERISTIQUES BIOPHYSIQUES DETERMINEES PAR SAXS POUR LE FRAGMENT TcCRTΔPΔ281 (PIC 3) | 140 |
| <u>TABLEAU 3.I.2</u> : QUANTITE MOYENNE DE PROTEINE OBTENUE EN FIN DE PURIFICATION POUR 1 L DE CULTURE BACTERIENNE | 141 |
| <u>TABLEAU 3.II.1</u> : STATISTIQUES DES DONNEES DE DIFFRACTION D'UN JEU NATIF (AVEC GLUCOSE) DE LA PROTEINE TcCRTΔPTERΔ288. Les valeurs entre parenthèse correspondent à la dernière tranche de résolution. | 151 |
| <u>TABLEAU 3.II.2</u> : STATISTIQUES DES DONNEES DE DIFFRACTION DES DEUX JEUX NATIFS DE LA PROTEINE EHCRTΔPΔ288. Les valeurs entre parenthèses correspondent à la dernière tranche de résolution | 154 |
| <u>TABLEAU 3.III.1</u> : STATISTIQUES D'AFFINEMENT DES STRUCTURES DES FRAGMENTS TcCRTΔPTERΔ288 ET EHCRTΔPΔ288, RESOLUES PAR REMPLACEMENT MOLECULAIRE. | 155 |
| <u>TABLEAU 3.IV.1</u> : CONSTANTES CINETIQUES DETERMINEES PAR SPR POUR L'INTERACTION GRC1Q-CALRETICULINES EN SUIVANT UN MODELE D'INTERACTION 1 :1 AVEC CHANGEMENT DE CONFORMATION | 191 |
| <u>TABLEAU 3.IV.2</u> : CONSTANTES CINETIQUES DETERMINEES PAR SPR POUR L'INTERACTION C1Q-CALRETICULINES EN SUIVANT UN MODELE D'INTERACTION 1 :1 AVEC CHANGEMENT DE CONFORMATION | 194 |
| <u>TABLEAU 3.VIII.1</u> : STATISTIQUES DES DONNEES DE DIFFRACTION DES JEUX NATIFS DES GRC1Q RECOMBINANTES. LES VALEURS ENTRE PARENTHESES CORRESPONDENT A LA DERNIERE TRANCHE DE RESOLUTION | 213 |
| <u>TABLEAU 3.VIII.2</u> : STATISTIQUES D'AFFINEMENT DE LA STRUCTURE DES GRC1Q RECOMBINANTES RESOLUE PAR REMPLACEMENT MOLECULAIRE | 213 |

ANNEXES

| | |
|--|---|
| <u>ANNEXE 4</u> : PARAMETRES PHYSICOCHIMIQUES DES PROTEINES UTILISEES AU COURS DE LA THESE | D |
|--|---|

Liste des abréviations

- ACAMP** : Apoptotic-cell associated molecular patterns
- ADN** : Acide Désoxyribonucléique
- ARN** : Acide Ribonucléique
- ARNm** : Acide Ribonucléique messenger
- B** : Facteur d'agitation thermique
- BCIP** : 5-bromo-4-chloro-indolyl-phosphate
- BSA** : Albumine de sérum bovin
- C1qR** : Récepteur de C1q
- CCP** : Module de type « Protéine de contrôle du complément » (Complement Control Protein)
- CCP4** : Collaborative Computational Project 4
- CLR** : Régions collagènes de C1q
- CMH** : Complexe Majeur d'Histocompatibilité
- CNX** : Calnexine
- CP1** : Protéase à cystéine 1
- CPBF** : Facteurs liés aux protéases à cystéine
- CRIT** : Inhibiteur du récepteur du complément C2
- CRP** : Protéine régulatrice du complément
- CRT** : Calréticuline
- CUB** : Module retrouvé dans complément C1r/C1s, Uegf, et Bmp1
- DFP** : Di-isopropylfluorophosphate
- Dmax** : Distance maximale
- DSS** : Acide 4,4diméthyl-4-silapentane-1-sulfonique
- DTT** : Dithiothréitol
- EDC** : 1-ethyl-3-(3-dimethylpropyl)-carbodiimide
- EDTA** : Acide éthylène diamine tétra-acétique
- EGF** : Module de type « facteur de croissance épidermique » (Epidermal Growth Factor)
- EhCRT** : Calréticuline de *Entamoeba histolytica*
- EhGL** : Lectine de *Entamoeba histolytica* spécifique du N-acetylglucosamine
- ELISA** : Enzyme linked Immunosorbent Assay
- ERAD** : Machinerie de dégradation associée au réticulum endoplasmique
- ERmanI** : Mannosidase I associée au réticulum endoplasmique
- ERO** : Espèces réactives de l'oxygène
- ESRF** : European Synchrotron Radiation Facility

GalNAc : N-acetyl galactosamine
gC1qR : Récepteur des régions globulaires de C1q
GH : Hormone de croissance
GI(II) : Glucosidase I(II)
GRC1q : Régions globulaires de C1q
HBS : Hepes-buffered saline
HEK : Human Embryonic Kidney
HRP : Horseradish peroxidase (Peroxydase de Raifort)
HSQC : Heteronuclear single quantum coherence spectroscopy
HuCRT : Calréticuline humaine
IBS : Institut de Biologie Structurale
IFN γ : Interféron γ
IgA, E, G, M : Immunoglobulines de type A, E, G ou M
IL : Interleukine
IMAC : immobilized metal ion affinity chromatography
IPTG : Isopropyl β -D-1-thiogalactopyranoside
LB : Milieu de culture Luria Bertani
LPS : Lipopolysaccharide
LRP1 : Low density lipoprotein receptor-related protein 1
LT : Lymphocyte T
MAC : Complexe d'attaque membranaire
MAE : Méthyl amino éthyl
MALDI-TOF : Matrix-Assisted Laser Desorption/Ionisation Time-Of_Flight (source de désorption/ionisation laser assistée par une matrice couplée à un analyseur à temps de vol)
MASP : MBL-Associated Serine Protease
MBL : Mannan-Binding Lectin (lectine reconnaissant le mannose)
MoCRT : Caléticuline de souris
NBT : Nitroblue tetrazolium chloride
NCS : Symétries non cristallographique
NGB : Domaine globulaire N terminal
NHS : N-hydroxysuccinimide
NSD : Normalized Standard Deviation (déviation standard normalisée)
p/v : Rapport poids/volume
PAMP : Pathogen-associated molecular patterns
PBS : Phosphate-buffered saline
PCR : Polymerase chain reaction
PDB : Protein Data Bank

PDI : Protein disulfide isomerase
PEG : Polyéthylène glycol
PS : Phostadidylsérine
PTX : Pentraxine
q : Fonction de momentum transer
RE : Réticulum endoplasmique
Req : Résonance à l'équilibre
Rg : Rayon de giration
Rmax : Résonance maximale
RMN : Résonance Magnétique Nucléaire
Rmsd : Root mean square deviation (déviation moyenne)
Rpm : rotations par minute
RU : Unité de résonance
SAXS : Diffusion des rayons X aux petits angles (Small Angles X-Ray Scattering)
SDS : Dodécylsulfate de sodium
SDS-PAGE : SDS-PolyAcrylamide Gel Electrophoresis (électrophorèse sur gel de polyacrylamide en présence de SDS)
SLE : Lupus érythémateux systémique
SHN : Sérum humain normal
SM : Spectrométrie de masse
SP : Protéase à sérine
SPR : Résonance plasmonique de surface
TAE : Tampon Tris Acétate EDTA
TCA : Trichloroacetic Acid (acide trichloroacétique)
TcCRT : Calréticuline de *Trypanosoma cruzi*
TEMED : Tétraméthyléthylènediamine
TLR : Récepteur de type Toll (Toll Like Receptor)
TLS : Translation/Libration/Screw
T_m : Température de fusion
TMB : 3,3'-5,5'-Tetramethylbenzidine
TNF α : Tumor Necrosis Factor α
Tris : Tris(hydroxyméthyl)-amino-méthane
TROSY : Transverse relaxation optimized spectroscopy
UGGT : UDP-glucose glycoprotéine transférase
v/v : Rapport volume/volume
WT : Souche sauvage (Wild type)

NOMENCLATURE UTILISEE POUR DESIGNER LES ACIDES AMINES

Les acides aminés ont été désignés en utilisant le code international des acides aminés à une et trois lettre selon la liste suivante

| <u>Acides aminés</u> | <u>Code 1 lettre</u> | <u>Code 3 lettres</u> |
|----------------------|----------------------|-----------------------|
| Acide aspartique | D | Asp |
| Acide glutamique | E | Glu |
| Alanine | A | Ala |
| Arginine | R | Arg |
| Asparagine | N | Asn |
| Cystéine | C | Cys |
| Glutamine | Q | Gln |
| Glycine | G | Gly |
| Histidine | H | His |
| Isoleucine | I | Ile |
| Leucine | L | Leu |
| Lysine | K | Lys |
| Méthionine | M | Met |
| Phénylalanine | F | Phe |
| Proline | P | Pro |
| Sérine | S | Ser |
| Thréonine | T | Thr |
| Tryptophane | W | Trp |
| Tyrosine | Y | Tyr |
| Valine | V | Val |

1. Introduction

I. La molécule C1q

I.1. Immunité innée et voie classique du complément

I.1.a. L'immunité innée

Pour lutter contre les agressions extérieures (bactéries, virus, blessures), notre organisme possède de nombreuses barrières physiques (épithélia, muqueuse pulmonaire, gastro-intestinale ou génitale) associées à une flore bactérienne protectrice. Une fois ces barrières franchies, il est nécessaire de limiter la pénétration des micro-organismes pathogènes dans la circulation sanguine et leur progression dans les tissus. A côté de ces menaces exogènes, le vieillissement des cellules peut entraîner des défauts dans les systèmes de réplication et de réparation de l'ADN pouvant nuire à la division cellulaire et entraîner la formation de cellules cancéreuses. Comme les micro-organismes pathogènes, ces cellules altérées doivent être éliminées par le système immunitaire.

La réponse immunitaire innée, qui est injustement considérée comme non spécifique, se met en place immédiatement à travers l'action des macrophages et des neutrophiles résidant dans les tissus et le sang, respectivement. Elle a pour effecteur essentiel la phagocytose, mécanisme central de la protection de l'organisme, qui conduit à l'internalisation de l'élément à éliminer par des cellules, puis à sa destruction.

Les phagocytes reconnaissent, via leurs récepteurs membranaires, des motifs particuliers à la surface des pathogènes (Pathogen Associated Molecular Patterns, PAMPs) : par exemple le mannose, le lipopolysaccharide (LPS), ou la LBP (Lipopolysaccharide binding protein), qui sont reconnus par les récepteurs CD206, CD14 et CD36, respectivement. Ils peuvent également reconnaître des motifs moléculaires associés au soi-altéré (Apoptotic Cell Associated Molecular Patterns, ACAMPs ; Danger Associated Molecular Patterns DAMPs).

La reconnaissance de motifs spécifiques à la surface des pathogènes entraîne le remodelage de la membrane et du cytosquelette pour « avaler » la particule à éliminer, qui sera contenue dans une vacuole issue de la membrane plasmique. A ce stade, la fusion entre la vacuole et un lysosome induit un changement radical des propriétés du lumen du phagolysosome nouvellement formé, assurant la dégradation de la particule ou cellule internalisée. En plus d'être très acide, il est également très riche en enzymes de dégradation (hydrolases, protéases, nucléases, liposidases, glycosidases), en peptides cationiques qui perméabilisent la

membrane, tandis que d'autres molécules comme la lactoferrine piègent des facteurs essentiels à la survie microbienne. Cette activité enzymatique de dégradation est couplée à la génération d'espèces réactives de l'oxygène (ERO) à travers l'action des NADPH oxydases présentes dans la membrane des macrophages (**Vieira et al., 2002**).

Même si la reconnaissance entre les cellules impliquées dans la phagocytose et les cibles à éliminer peut impliquer une interaction directe, l'efficacité de cette dernière peut être grandement améliorée par l'intermédiaire de molécules nommées opsonines (comme les immunoglobulines ou le produit de la cascade du complément C3b), qui recouvrent la surface de la cible en fournissant des « signaux » supplémentaires aux phagocytes. Il a été montré en effet que les macrophages possédaient à leur surface de nombreux récepteurs aux immunoglobulines (CD64, CD32a, CD89, CD16) et aux fragments de la protéine C3 (CR1, CR3, CR4) (**Jaumouillé and Grinstein, 2011; Ricklin et al., 2010**).

I.1.b. C1q dans la voie classique du système du complément

En 1900, Jules Bordet a identifié un composant thermosensible présent dans le sérum et nécessaire à la lyse des bactéries et des érythrocytes, déclenché par les complexes immuns. Cette substance, tout d'abord nommée alexine, est en fait composée d'une trentaine de protéines présentes à une concentration de trois grammes par litre de plasma, d'où le nom de système. Même si ce système a initialement été proposé comme un « complément » de l'activité des anticorps, il tient en réalité une place importante et majeure au sein de l'immunité innée en identifiant et signalant les dangers, recrutant les effecteurs de l'immunité innée sur le lieu de l'infection, en participant à la lyse des pathogènes et à la maturation de certains acteurs de l'immunité adaptative.

La finalité de la cascade du complément est de générer en grande quantité le fragment protéique C3b jouant le rôle d'opsonine et de participer à la mise en place du complexe d'attaque membranaire (MAC). De plus, elle génère une série d'anaphylatoxines permettant de moduler la réponse cellulaire et l'immunité adaptative et d'attirer par chimiotactisme les cellules responsables de la phagocytose. Même si la cascade du complément est unique, elle est initiée par trois voies d'activation distinctes.

- **La voie « classique »** est initiée par la protéine C1q associée aux protéases C1r et C1s que je décrirai à un niveau structural plus précis dans la prochaine section. Cette

activation est principalement déclenchée suite à la reconnaissance d'opsonines IgM, IgG et pentraxines (comme CRP), fixées à la surface des pathogènes.

- **La voie « lectine »** est déclenchée par d'autres protéines similaires à C1q, les ficolines H, L, M et les collectines MBL et CL-LK associées aux protéases MASPs (1-2-3), similaires à C1r et C1s. Les collectines possèdent un domaine lectine de type C qui permet la reconnaissance des monosaccharides via leurs groupements 3'-OH et 4'-OH et des N-aminoglycanes couramment présents à la surface des bactéries et virus et absents des cellules de l'hôte.
- **La voie alterne** est initiée par l'hydrolyse spontanée de la protéine C3 en fragments C3a et C3b.
- **Une quatrième voie** associée aux mécanismes de coagulation pourrait également exister en faisant intervenir des protéases de la cascade de la coagulation (plasmine, thrombine, élastase) (**Amara et al., 2008**).

La fixation de C1q induit un changement de conformation de la protéine permettant la transduction du signal depuis les régions globulaires jusqu'aux régions collagènes associées aux protéases, en provoquant une activation mécanique des protéases C1r-C1s (**Gaboriaud et al., 2004**), que je décrirai à un niveau structural dans la section suivante (**Figure 1.I.1, A**).

La protéase C1s, une fois activée, clive la protéine C4 en fragments C4a et C4b. C4b peut alors se fixer à la surface du pathogène à éliminer par des liaisons covalentes de type thioester et recruter C2, qui est alors hydrolysé en fragments C2a et C2b par C1s, pour former la C3 convertase (C4b2a), dont le rôle est de cliver la protéine C3 en fragments C3a et C3b (**Figure 1.I.1, B**). Comme C4b, C3b expose un groupe thioester très réactif, lui permettant de se fixer covalamment à la membrane du pathogène. Lorsque l'hydrolyse des différents fragments s'effectue, seul ceux à proximité de la membrane se fixent, car la liaison thioester très instable, peut réagir avec les molécules d'eau pour former un fragment inerte C3b(H₂O) (**Krishnan et al., 2009**).

L'hydrolyse spontanée des molécules de C3 en C3H₂O (similaire à C3b) déclenche la voie alterne (**Lachmann and Halbwachs, 1975; Pangburn et al., 1981**). C3H₂O peut alors s'associer au facteur B pour former le complexe C3H₂OB, qui est immédiatement clivé par le facteur D pour former la C3 convertase alterne, capable de cliver de nouvelles molécules de

C3 et d'amplifier le signal (**Figure 1.I.1, F**) (**Isenman et al., 1981; Li et al., 2010a; Nishida et al., 2006; Pangburn et al., 1981; Rodriguez et al., 2014; Winters et al., 2005**). Même si cette voie est considérée comme non spécifique, la properdine, protéine stabilisant le complexe C3 convertase (**Fearon and Austen, 1975**), reconnaît des motifs particuliers à la surface des cellules ou des pathogènes comme l'héparine, ce qui permet de diriger la voie alterne vers des cibles spécifiques (**Zaferani et al., 2011**). La voie alterne joue également le rôle d'une boucle d'amplification des voies classique et lectine du complément.

La production en grande quantité du fragment C3b peut conduire ce dernier à se fixer covalamment à la C3 convertase pour former la C5 convertase (C4b2a3b) (**Figure 1.I.1, C**). Cette dernière peut alors interagir avec la protéine C5 et l'hydrolyser en fragments C5a et C5b (**Figure 1.I.1.b, D**), (**Laursen et al., 2011**). Il a été montré également que C5 pouvait être hydrolysé spontanément par des facteurs impliqués dans la coagulation comme la thrombine (**Amara et al., 2010**).

Le clivage de C5 en C5b induit un changement de conformation lui permettant de se fixer à la protéine C6 par le domaine TED (domaine thioester). Les protéines C7, C8 et au moins 18 copies de la protéine C9 se rajoutent à cette plateforme, aboutissant à la formation du complexe d'attaque membranaire MAC (**Preissner et al., 1985**). Dans ce complexe, la protéine C8 pénètre dans la membrane plasmique grâce à une structure similaire à celle des perforines (**Hadders et al., 2007**) et les protéines C9 forment le canal à travers la membrane (**Figure 1.I.1, E**) (**Aleshin et al., 2012a, 2012b; Bhakdi and Trantum-Jensen, 1991; Bramham et al., 2005; Laursen et al., 2012; Lovelace et al., 2008, 2011; Rawal and Pangburn, 2001; Slade et al., 2008**).

Parmi les fragments générés lors de la cascade du complément, les anaphylatoxines C3a et C5a jouent un rôle important dans l'inflammation et la maturation du système immunitaire adaptatif. Elles recrutent les macrophages, éosinophiles et neutrophiles sur le lieu de l'infection en se fixant à des récepteurs spécifiques des anaphylatoxines (**Aksamit et al., 1981; Ehrenguber et al., 1994; Elsner et al., 1994a, 1994b; Merle et al., 2015a; Murakami et al., 1993**).

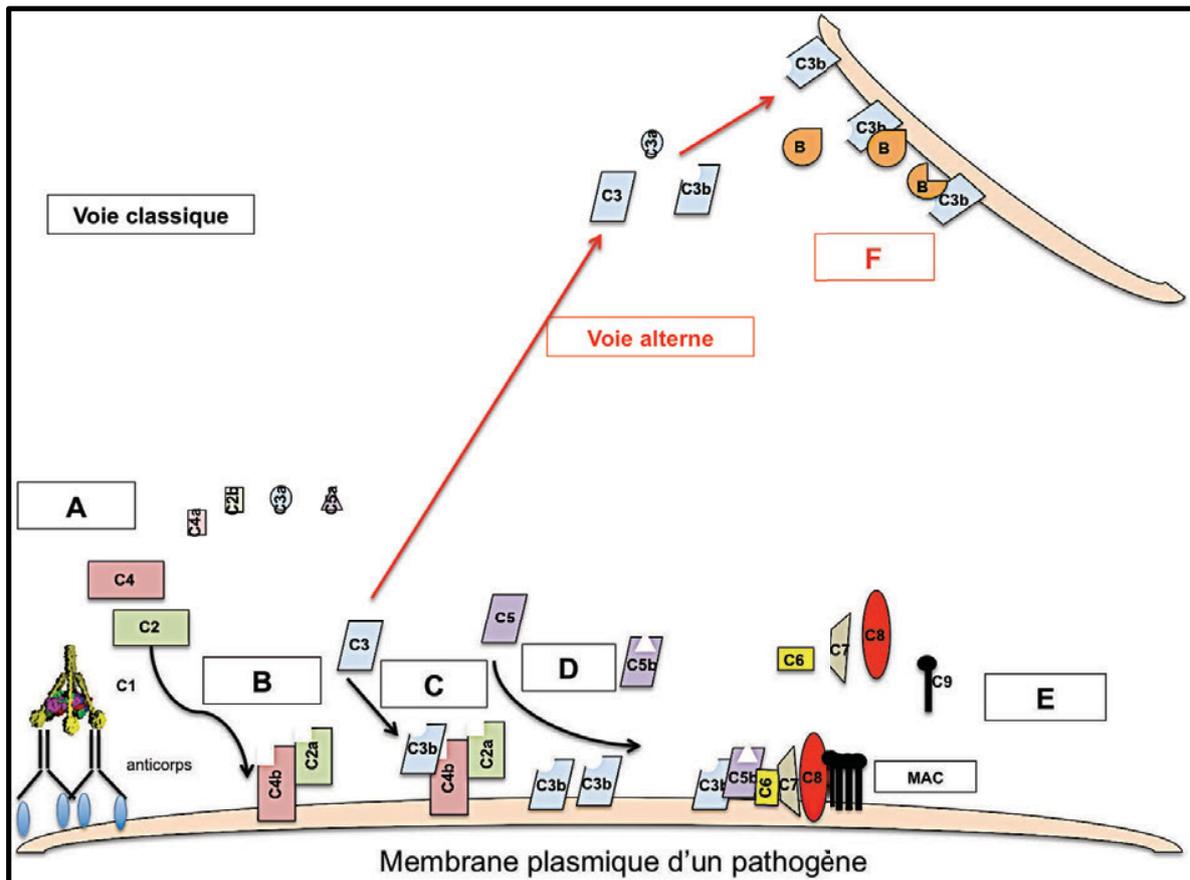


FIGURE 1.1.1: REPRÉSENTATION SCHEMATIQUE DE LA CASCADE DU COMPLÉMENT, ACTIVÉE PAR LES VOIES CLASSIQUE ET ALTERNE

L'importance de la réponse inflammatoire générée et du pouvoir de dégradation de la cascade du complément et son amplification par la voie alterne peuvent vite en faire une voie incontrôlable et destructrice, qui doit absolument être régulée par des médiateurs solubles et membranaires. Par exemple, C1-INH (inhibiteur de protéase-serpine) inhibe la cascade du complément à un stade précoce en se fixant covalamment aux protéases C1r et C1s et les décrochant de C1q suite à un changement conformationnel (idem pour les protéases MASPs). Tandis que d'autres effecteurs comme C4BP (en ciblant C4 et C3) et les facteurs H, CFHL-1 et MCP (en ciblant C3b) jouent le rôle de co-facteur pour le facteur I et sont impliqués dans la dissociation de la C3 convertase et l'inactivation de C3b et 4b en les clivant en iC3b, C3c, C3dg ou C4c, C4d respectivement. Plus tardivement, des effecteurs membranaires (CD59 empêchant l'oligomérisation de C9) et plasmatiques (vitronectine et clusterine en bloquant la fixation et C5b-C7 et de C8 à la membrane, respectivement) agissent également pour freiner la formation du MAC, qui empêchent l'insertion des composants du MAC dans la membrane. Tandis que DAF et CR1 dissocient les C3 et C5 convertases (Carroll and Sim, 2011; Ricklin et al., 2010).

I.2. Aspects structuraux de C1q, ses domaines de reconnaissance et de l'activation du complexe C1 (C1q-C1r₂-C1s₂)

I.2.a. C1q

C1q est une protéine multimérique composée de 18 chaînes polypeptidiques de trois types différents : A, B et C, comportant respectivement 223, 236 et 217 résidus, présentant une masse moléculaire estimée à 459300 Da (Reid, 1983) et depuis confirmée à 460793 Da par spectrométrie de masse (MS) (Reid, 1974, 1982; Sellar et al., 1991; Tissot et al., 2005). Chacune de ces chaînes possède une courte région N-terminale (3 à 9 résidus) contenant des ponts disulfures, suivie par des régions de type collagène (CLR) d'environ 81 résidus et un domaine globulaire C-terminal d'environ 135 résidus (GRC1q). Les chaînes sont associées par groupe de trois (ABC) par des interactions non covalentes, formant des hélices de type collagène, suivies en C-terminal par les domaines globulaires. L'assemblage est consolidé par un pont disulfure intermoléculaire entre les chaînes A et B, et un autre entre deux chaînes C. C1q apparaît sous sa forme caractéristique de bouquet de tulipes en microscopie électronique (Diebolder et al., 2014; Kishore and Reid, 1999; Kishore et al., 2002; Knobel et al., 1975) (Figure 1.I.2, A et B).

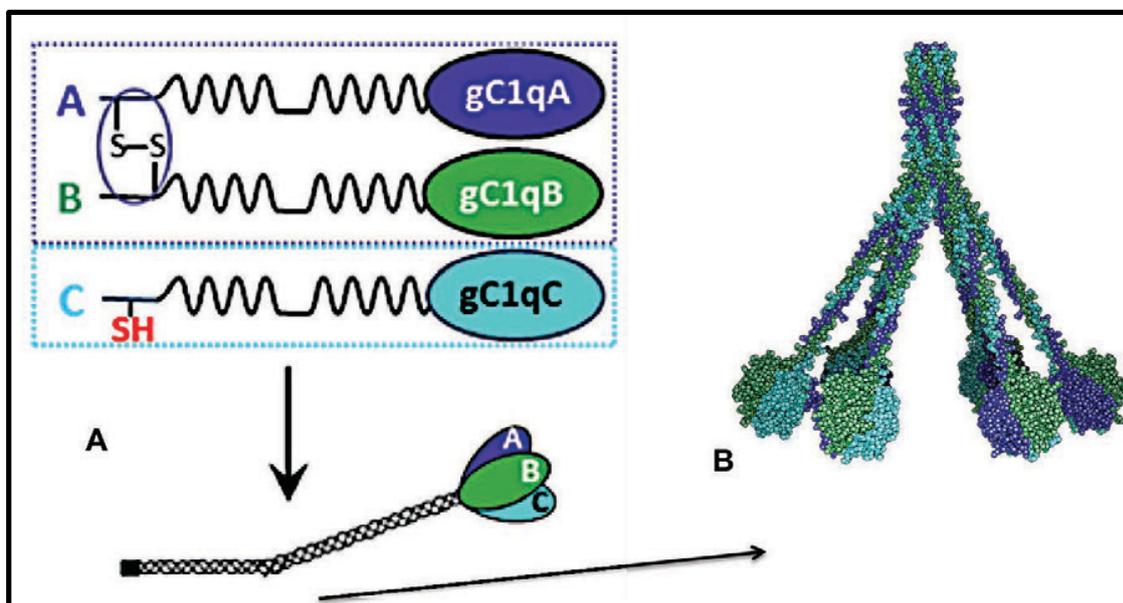


FIGURE 1.I.2: ORGANISATION STRUCTURALE DE C1Q (

(A) Trois chaînes A, B et C s'associent par des liaisons non covalentes et un pont disulfure A-B pour former une unité structurale. **(B)** Le groupement thiolate libre sur la chaîne C permet à deux unités structurales de s'assembler via un pont disulfure C-C pour former une unité composée de 6 chaînes dans l'ordre ABC-CBA. L'ensemble de la molécule C1q obtenue par modélisation moléculaire à partir de la structure des régions globulaires présente une forme de 'bouquet de tulipes' dont les tiges sont constituées des hélices de type collagène.

Les triples hélices de type collagène sont constituées par la répétition de séquences G-X-Y. L'interruption de cette répétition de séquences à la mi-longueur, avec l'insertion d'une thréonine dans la chaîne A et la substitution d'une alanine en glycine dans la chaîne C, provoque la divergence des 6 tiges du bouquet. En plus de produire cette discontinuité, ces modifications fournissent une importante flexibilité à la molécule (**Hanson et al., 1985; Poon et al., 1983**).

La protéine C1q possède de nombreuses modifications post-traductionnelles situées principalement au niveau des lysines et prolines des CLR (seules les 3 lysines K59^A, K61^B et K58^C ne sont pas modifiées) (**Pflieger et al., 2010**). La digestion de C1q par la collagénase permet d'isoler les GRC1q, alors que la protéolyse limitée par la pepsine à pH 4,5 permet d'isoler les CLR. Ces protocoles ont permis d'étudier séparément l'activité de ces domaines.

Ainsi la reconnaissance et la fixation des ligands ont largement été attribuées aux régions globulaires (**Hughes-Jones and Gardner, 1979; Kishore et al., 2002; Tschopp et al., 1980**) alors que les CLR sont majoritairement responsables de l'interaction avec les protéases C1r-C1s, qui portent l'activité catalytique (**Bally et al., 2013; Reid et al., 1977**), même s'il a été montré qu'ils pouvaient également interagir avec certains ligands ou récepteurs (**Jiang et al., 1992, 1994; Johns and Bernard, 1997**).

I.2.b. Structure cristallographique des GRC1q

Les propriétés de reconnaissance des régions globulaires de C1q ont pu être décryptées grâce à la détermination de la structure par cristallographie aux rayons X à une résolution de 1.3 Å (**Gaboriaud et al., 2003; Garlatti et al., 2010**) (PDB : 1PK6 et 2WNV). Les régions globulaires forment une structure hétérotrimérique compacte, presque sphérique, avec un diamètre de 50 Å et un axe de pseudo-symétrie non cristallographique d'ordre 3.

La structure montre l'émergence des extrémités N- et C-terminales à la base du trimère et un repliement de type « jelly-roll » organisé autour d'un sandwich de 10 brins β et 5 brins β anti parallèles. Les trois sous-unités A, B et C sont maintenues entre elles par des interactions latérales hydrophobes à la base et, près de l'apex le long de l'axe 3, des interactions polaires et hydrophiles (**Figure 1.I.3, A et B**).

Au niveau de l'apex, un ion calcium est coordonné par six atomes d'oxygènes (3 des chaînes latérales des résidus D172^B, Q177^A, Q179^B, un oxygène de la chaîne principale de Y173^B et deux molécules d'eau) (**Figure 1.I.3, C**).

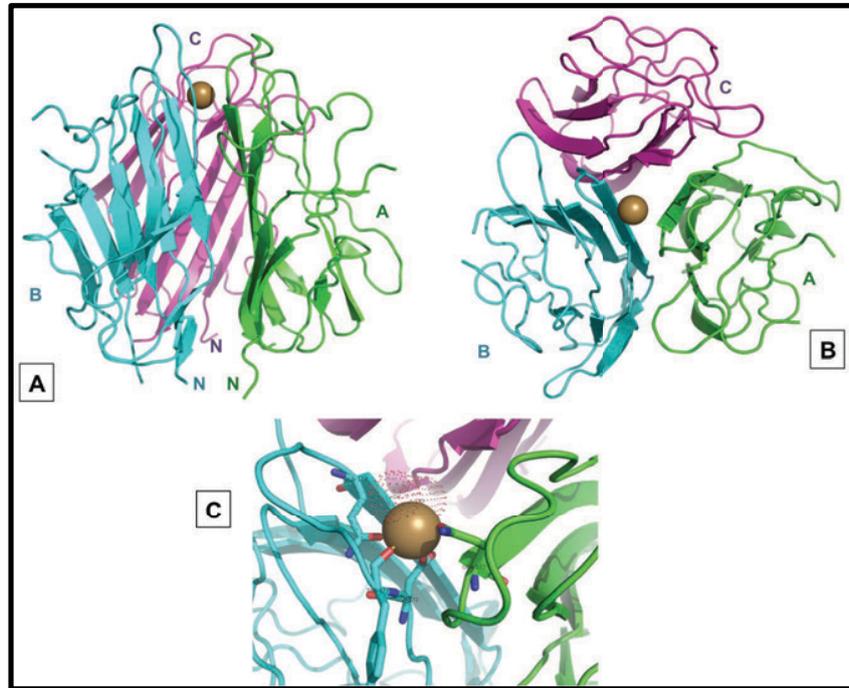


FIGURE 1.I.3: STRUCTURE CRISTALLOGRAPHIQUE DES GRC1Q (ADAPTE DE (GABORIAUD ET AL., 2003))

L'organisation des trois chaînes A, C, B formant le domaine globulaire est représentée de côté (A) et vue du dessus (B). La figure C représente les acides aminés impliqués dans la fixation de l'atome de calcium à l'apex.

La distribution des charges à la surface des chaînes A, B et C n'est pas équivalente. A et C présentent des résidus basiques (R pour la chaîne A) et acides sur la surface externe du domaine et quelques résidus lysine au sommet. La chaîne B expose des charges positives de manière prépondérante, spécifiquement un cluster de résidus arginine à la surface latérale (R101^B, R114^B, R179^B) et à la jonction des surfaces apicales et latérales (R108^B, R109^B).

La prédominance de charges positives à sa surface et la position équatoriale de la chaîne B la rendent très accessible aux interactions avec des cibles chargées négativement (**Gaboriaud et al., 2003, 2011a; Garlatti et al., 2010**).

I.2.c. Reconnaissance des immunoglobulines

C1q est capable de reconnaître une très grande variété de ligands, mais l'identification précise des acides aminés impliqués dans ces interactions reste un challenge. Les sites de fixation de C1q étant majoritairement chargés, il est proposé que la protéine soit un « senseur » de charges plutôt qu'impliquée dans des interactions spécifiques de type enzyme-substrat (**Ghai et al., 2007; Kishore et al., 2002**).

En effet, les premières caractérisations de l'interaction entre C1q, les IgG et les IgM ont confirmé la contribution importante des interactions ioniques (**Agrawal et al., 2001; Alberti et al., 1996; Burton et al., 1980; Comis and Easterbrook-Smith, 1985; Easterbrook-Smith, 1983; Groeneveld et al., 2005; Poon et al., 1983; Rhodes, 2000; Wines and Easterbrook-Smith, 1990**). Des expériences de réticulation par pontage chimique ont mis en évidence l'implication des trois chaînes A, B et C dans les interactions avec les IgG (**Wines and Easterbrook-Smith, 1990**). Un premier site de fixation a été suggéré sur la chaîne B, incluant les deux arginines 114 et 129, alors qu'un second inclut les trois chaînes et les résidus R162^A, R163^B, R156^C (**Marqués et al., 1993**), tandis que d'autres études suggèrent indépendamment un site proche du résidu R109^B (**Tissot et al., 2005**). Dans les quinze dernières années, des analyses de modélisation et de mutations (**Gaboriaud et al., 2003; Kojouharova et al., 2004; Schneider and Zacharias, 2012; Zlatarova et al., 2006**) ont suggéré la participation des résidus R109^B, R114^B, R129^B et E162^B dans la fixation des IgG. Le premier modèle liant la structure cristallographique avec des approches complémentaires comme la caractérisation physicochimique des interactions ioniques, les modifications chimiques (**Burton et al., 1980; Marqués et al., 1993; McCall and Easterbrook-Smith, 1989**) et l'analyse de mutations des IgG, a permis de proposer un modèle d'interaction dans lequel les résidus R129^B and E162^B forment des interactions ioniques avec D270 et K322 de l'IgG au niveau du domaine CH₂, tandis que les résidus R114^B et R161^B fournissent des interactions ioniques supplémentaires (**Gaboriaud et al., 2003**) (**Figure 1.I.4**).

Le rôle de l'ion calcium dans les interactions est quant à lui sujet à discussion et semble dépendant du type de ligand. La modélisation *in silico* montre que la perte de l'ion calcium modifie la direction du moment dipolaire électrique d'une position axiale centrée sur l'apex à une position latérale centrée sur la chaîne B (**Roumenina et al., 2005**), deux faces contenant les résidus considérés comme importants dans les différentes interactions (**Gaboriaud et al., 2003; Ghai et al., 2007; Kishore et al., 2004**). Il a été montré que la présence du calcium

influence les interactions avec les IgM, IgG, CRP, PTX3 et LPS (Roumenina et al., 2005, 2006, 2008).

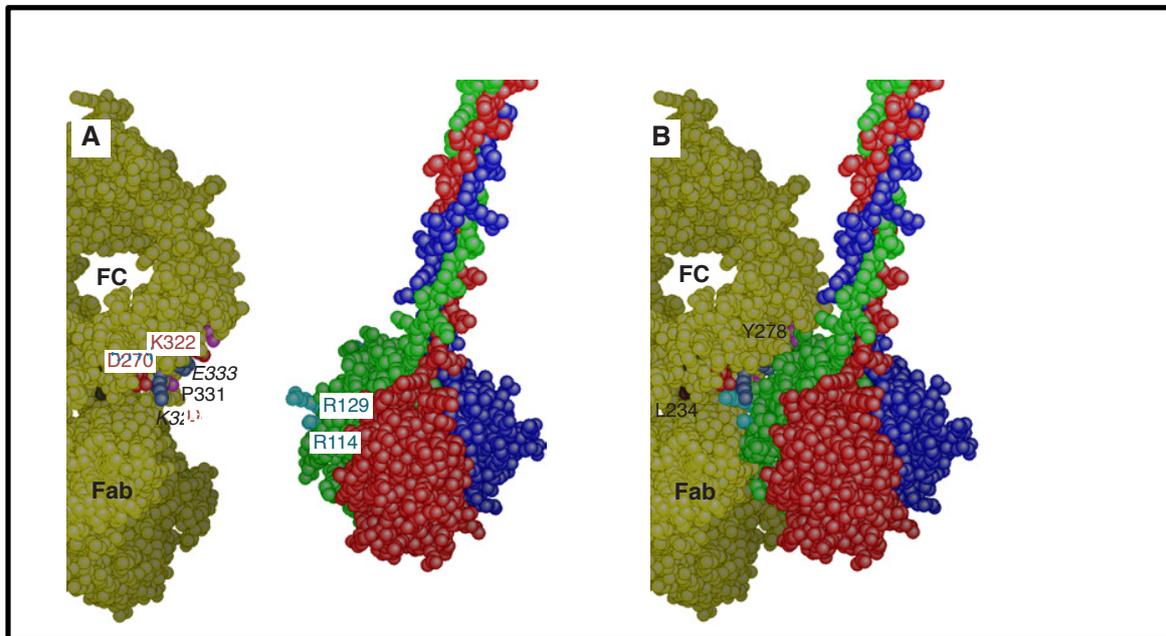


FIGURE 1.I.4: MODELISATION DE LA SURFACE D'INTERACTION ENTRE LES GRC1Q ET LES IGG (GABORIAUD ET AL., 2003)

I.2.d Activation des protéases C1r_s-C1s₂

Les deux protéases associées à C1q dans le complexe C1, C1r (90 kDa) et C1s (80 kDa) sont présentes initialement sous la forme zymogène. Elles possèdent deux domaines CUB (module retrouvé dans complément C1r/C1s, Uegf, et Bmp1) séparés par un domaine EGF (epidermal growth factor), suivis de deux modules de contrôle du complément (CCP) et un domaine C-terminal de type protéase à sérine (SP) (Gaboriaud et al., 2014) (Figure 1.I.5). En absence de C1q, les protéases ont une forme allongée en S avec les deux protéines C1r au centre, assemblées par leur domaine catalytique et stabilisées en conformation tête-bêche, le domaine CCP1 de l'un des monomères interagissant avec le domaine SP de l'autre. Chaque monomère de C1r est associé à un monomère de C1s par une interaction dépendante du calcium, via un assemblage tête-bêche des domaines CUB1 et EGF (Budayova-Spano et al., 2002a; Gregory et al., 2003; Tschopp et al., 1980) (Figures 1.I.6 et 1.I.7).

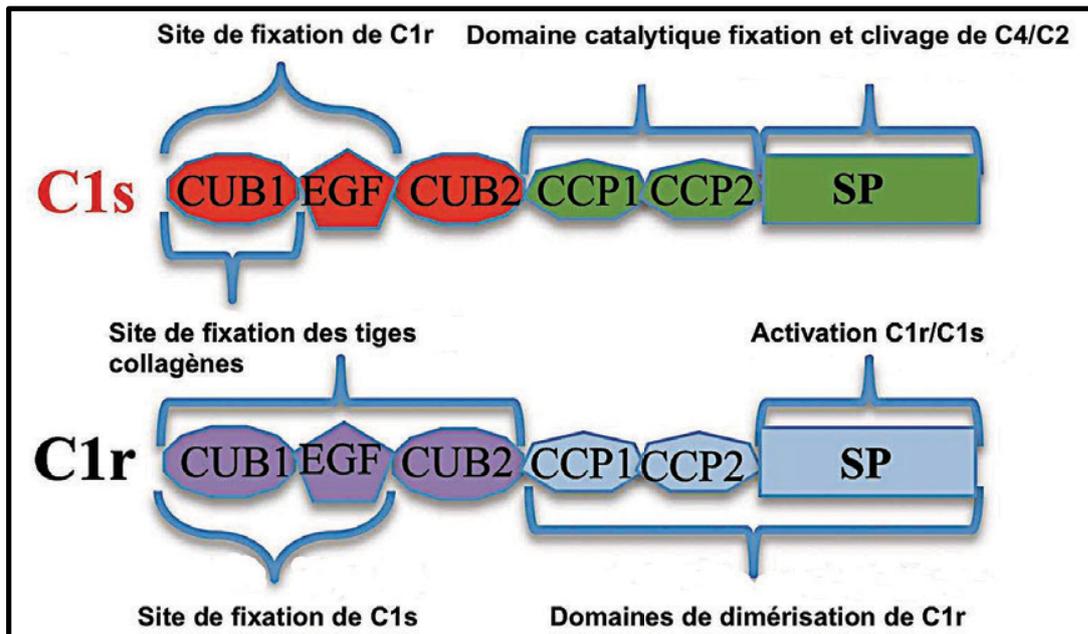


FIGURE 1.I.5 : ORGANISATION MODULAIRE DES PROTEASES C1R ET C1S (GABORIAUD ET AL., 2014)

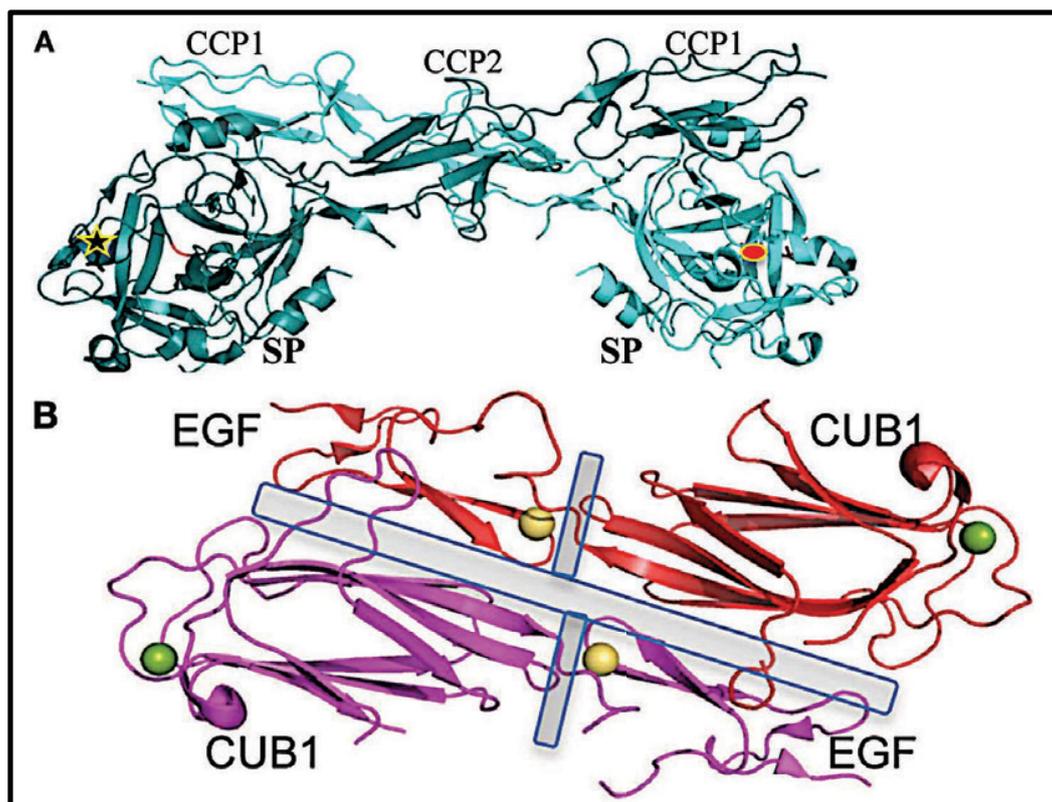


FIGURE 1.I.6 : DIMERISATION DES PROTEINES C1R ET C1S (GABORIAUD ET AL., 2014)

(A) La résolution de la structure des domaines CCP1-CCP2-SP de C1r montre l'assemblage du dimère. (B) L'interaction dans le cristal d'un homodimère de C1s dans une conformation tête-bêche via l'interaction entre les domaines CUB et EGF suggère le même type d'association pour l'hétérodimère C1r-C1s.

De manière surprenante, la protéine C1q et le complexe C1 possèdent un rayon de giration équivalent (environ 12,6-12,8 nm), alors que les protéases seules en solution présentent un rayon de 17 nm, mesuré par diffusion des neutrons, ce qui suggère un changement de conformation lors de leur association à C1q (**Figure 1.I.7, A**) (**Arlaud et al., 1987; Perkins et al., 1984**). Des mutations introduites dans les protéines C1s, C1r et plus récemment C1q (grâce à la production récente de la première forme recombinante de C1q) ont permis de localiser précisément les sites d'interactions entre les protéases et C1q. Ces derniers impliquent les résidus E49/Y56/D102 du domaine CUB₁ et D226/D228/Y235/D273 du CUB₂ de C1r et les résidus E45 et D98 du domaine CUB₁ de C1s (**Bally et al., 2009; Thielens et al., 1994; Tsai et al., 1997**), ainsi que les lysines 61^B et 58^C et la glycine 63^B côté C1q (**Bally et al., 2013; Roumenina et al., 2011**). Grâce à ces expériences, et les différentes structures obtenues pour les différents modules de C1r et C1s, un modèle du complexe C1 a pu être proposé, dans lequel un ion calcium participe à l'interaction de C1s et C1r avec les collagènes (**Bally et al., 2009; Gaboriaud et al., 2011b; Venkatraman Girija et al., 2013**).

Dans le complexe C1, les protéases doivent être activées par un clivage protéolytique spécifique entre des résidus arginine et isoleucine. Les différentes conformations observées lors de la résolution des structures obtenues par cristallographie aux rayons X des domaines SP (**Budayova-Spano et al., 2002a; Perry et al., 2013**) ont permis d'apporter des informations sur la spécificité de substrat et les mécanismes d'activation (**Budayova-Spano et al., 2002a; Gaboriaud et al., 2000; Perry et al., 2013**). Cependant la position du site catalytique de C1s est encore controversée. En effet alors que les données structurales précédemment citées le positionne à l'extérieur du cône formé par les tiges collagènes de C1q, des recherches plus récentes le situent à l'intérieur (**Venkatraman Girija et al., 2013**). Ce dernier modèle implique néanmoins une pénétration peu probable des molécules C4 et C2 à l'intérieur du cône. Il a également été montré que le domaine SP seul ne suffit pas pour assurer la coupure de C4 par la protéase C1s, suggérant une participation des domaines CCP (**Rossi et al., 1998**). Les protéases enfouies dans le cône de C1q ont une forme de huit, et la distance de 90 Å entre les sites catalytiques et d'activation de C1r préserve la protéase d'une auto activation (**Figure 1.I.7, B et C**).

Le mécanisme par lequel cette activation se produit est encore mal connu. Cependant l'ion calcium situé au niveau de l'apex des GRC1q pourrait jouer un rôle important. Dans ce nouveau modèle, la fixation d'un ligand modifierait localement la conformation des GRC1q et libérerait l'ion calcium, entraînant une modification du moment dipolaire, et

éventuellement une rotation des régions globulaires. Cette rotation pourrait entraîner un stress mécanique à l'origine de l'ouverture des tiges collagènes, elle-même nécessaire à la réorganisation des protéases au cours de l'étape d'activation (Arlaud et al., 2001, 2002a, 2002b; Budayova-Spano et al., 2002a, 2002b; Gaboriaud et al., 2004; Roumenina et al., 2005).

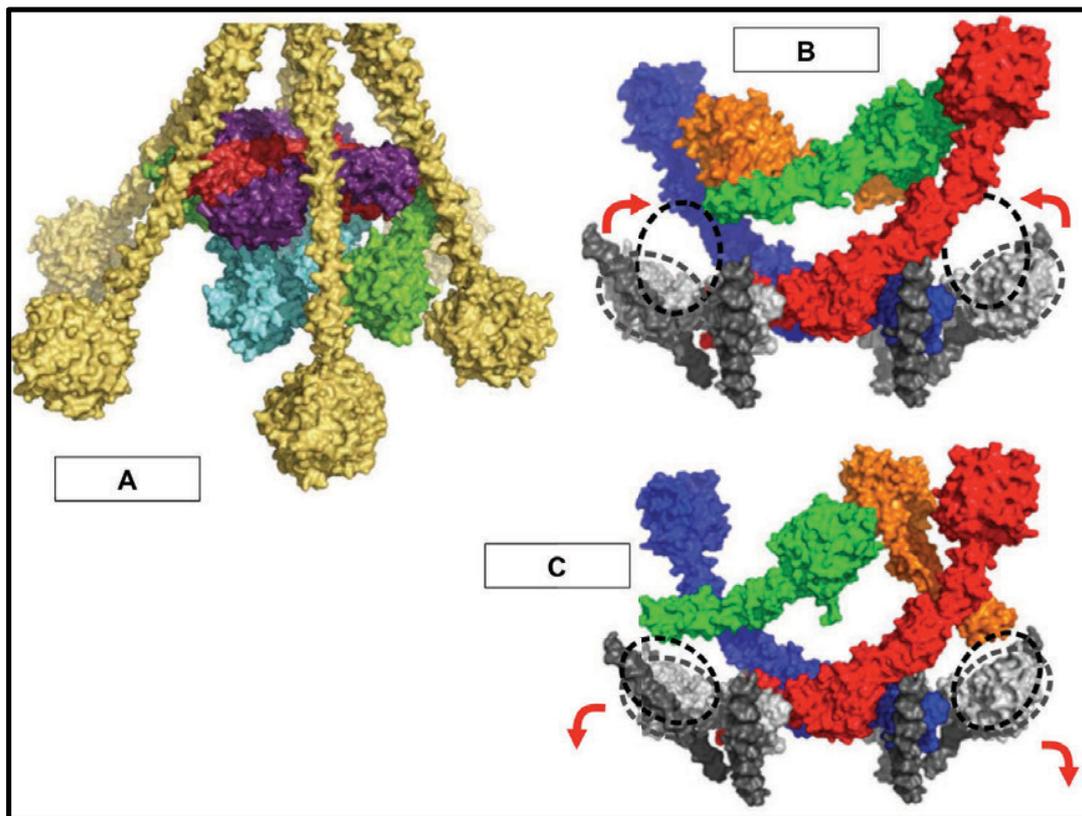


FIGURE 1.I.7: MODELISATION DU POSITIONNEMENT DES PROTEASES C1R ET C1S DANS LE COMPLEXE C1 (GABORIAUD ET AL., 2014; VENKATRAMAN GIRIJA ET AL., 2013)

(A) Les protéases C1r et C1s seraient compactées à l'intérieur du cône de C1q, contrairement à leur forme allongée en solution. (B) Dans la forme inactive, les sites d'activation et catalytiques du dimère de C1r représenté en orange et vert sont éloignés. (C) L'activation mécanique du complexe C1 rapproche les sites catalytiques et d'activation des protéases C1r qui, une fois activées, sont capables de cliver les protéases C1s représentées en bleu et rouge.

I.3. Rôle de C1q dans l'efferocytose

Au laboratoire, nous nous intéressons à l'efferocytose (venant littéralement du mot latin efferre, 'conduire à la tombe'), c'est-à-dire au processus fondamental de l'élimination rapide et non inflammatoire par phagocytose (processus cellulaire permettant à une cellule d'engouffrer et de dégrader une autre cellule ou une particule étrangère) des cellules en fin de

vie ou soumises à un stress important. L'apoptose est un processus naturel qui permet d'éliminer de manière programmée des cellules en fin de vie, ce qui est essentiel pour assurer la maintenance d'un organisme. Plusieurs voies métaboliques interconnectées conduisent à la mort de la cellule qui sera éliminée. Je décrirai ici comment C1q participe à ce processus.

I.3.a. Rôle de C1q dans la phagocytose

C1q améliore la phagocytose de manière indépendante de la voie classique du complément (**Bohlsón et al., 2007**). En effet C1q est capable de se lier directement aux cellules impliquées dans la phagocytose et la réponse immunitaire, comme les LB, monocytes, macrophages, DCs, neutrophiles en utilisant de nombreux récepteurs incluant cC1qR, CD91, CR1, CD93 et gC1qR (**Galvan et al., 2012**) et améliore l'efficacité de ces cellules (**Bobak et al., 1987, 1988**).

Dans ce contexte, C1q interagirait avec les cellules par l'intermédiaire de ses domaines collagènes, décrits comme essentiels à l'amélioration de la phagocytose. Même si les mécanismes de cette fonction de C1q et les récepteurs spécifiques sont encore méconnus, il se pourrait que C1q serve de molécule pontante entre les phagocytes et les cellules apoptotiques, en interagissant principalement avec le complexe calréticuline-CD91 (**Ogden et al., 2001**). Cependant la suppression de CD91 chez la souris n'entraîne pas clairement la diminution de l'activité de C1q dans l'amélioration de la phagocytose (**Galvan et al., 2012**).

I.3.b. Implication de C1q dans le SLE et l'élimination des cellules apoptotiques

Etonnamment, alors que l'on peut s'attendre à ce qu'une délétion ou une mutation de C1q entraîne pour l'organisme uniquement une sensibilité accrue aux infections et une diminution de la réponse immunitaire, dans 90% des cas elle entraîne également l'apparition d'une maladie auto-immune : le lupus érythémateux systémique (SLE), chez l'homme comme chez la souris (**Martens et al., 2009; Pickering et al., 2000**). Les patients atteints par le SLE présentent une accumulation de cellules apoptotiques, qui ne sont plus éliminées par les macrophages (**Botto et al., 1998; Taylor et al., 2000**). Cet accroissement de cellules apoptotiques est la source de nouveaux antigènes et permet le développement d'auto-anticorps dirigés contre des éléments du soi, comme l'ADN ou des phospholipides (**D'Cruz et al., 2007; Muñoz et al., 2010a, 2010b**).

Ces observations ont mis en évidence le rôle de C1q dans l'élimination des cellules apoptotiques. Cependant une question reste en suspens: est-ce que C1q intervient aussi dans le contexte de la cascade du complément, en participant au complexe C1 ou seulement indépendamment des protéases C1r-C1s ? C1q peut induire le dépôt de C3b à la surface des cellules apoptotiques (**Donnelly et al., 2006; Fraser et al., 2009; Nauta et al., 2002; Quartier et al., 2005; Taylor et al., 2000**), de plus il a été montré que la présence des IgM et des IgG dirigés contre les cellules apoptotiques jouait un rôle important dans leur phagocytose et leur élimination, via les récepteurs CR3 et CR4 (**Mevorach et al., 1998; Quartier et al., 2005; Takizawa et al., 1996**). Ces résultats suggèrent une participation de C1q dépendante du complément dans l'élimination des cellules apoptotiques. Néanmoins, l'absence d'inflammation peut laisser entrevoir d'autres types d'intervention. En effet, même si les cellules humaines possèdent de nombreux inhibiteurs du complément à leur surface permettant de réguler l'inflammation, il se pourrait que C1q participe à leur élimination sans être lié aux protéases C1r et C1s.

I.3.c. Reconnaissance par C1q de motifs moléculaires à la surface des cellules apoptotiques.

C1q est capable de reconnaître différents motifs à la surface des cellules apoptotiques nommés ACAMPs par l'intermédiaire de ses régions globulaires (**Gaboriaud et al., 2011a; Nauta et al., 2002**). Parmi ces différents motifs reconnus, de nombreux candidats ont été identifiés comme jouant le rôle de signaux « eat-me » en favorisant la phagocytose des cellules apoptotiques *via* l'interaction avec C1q.

C'est notamment le cas de la phosphatidylsérine (**Païdassi et al., 2008a, 2011**) ; l'ADN double brin (**Païdassi et al., 2008b; Tissot et al., 2003**), la GAPDH (**Terrasse et al., 2012**), les annexines A2 et A5 (**Martin et al., 2012**), mais également la calréticuline (**Ogden et al., 2001; Païdassi et al., 2011; Verneret et al., 2014**). L'interaction de C1q avec ces signaux « eat-me » permettrait de favoriser l'interaction des cellules avec les phagocytes et la phagocytose, indépendamment de la voie classique du système du complément (**Fraser et al., 2010; Nauta et al., 2004; Ogden et al., 2001**).

L'analyse de la structure des régions globulaires de C1q (GRC1q) à haute résolution en complexe avec le désoxyribose pour l'ADN et la phospho-sérine (PS) a permis de proposer un mécanisme impliquant une localisation différente des sites de liaison selon que l'interaction entraîne une réponse inflammatoire ou non (**Figure 1.I.8**).

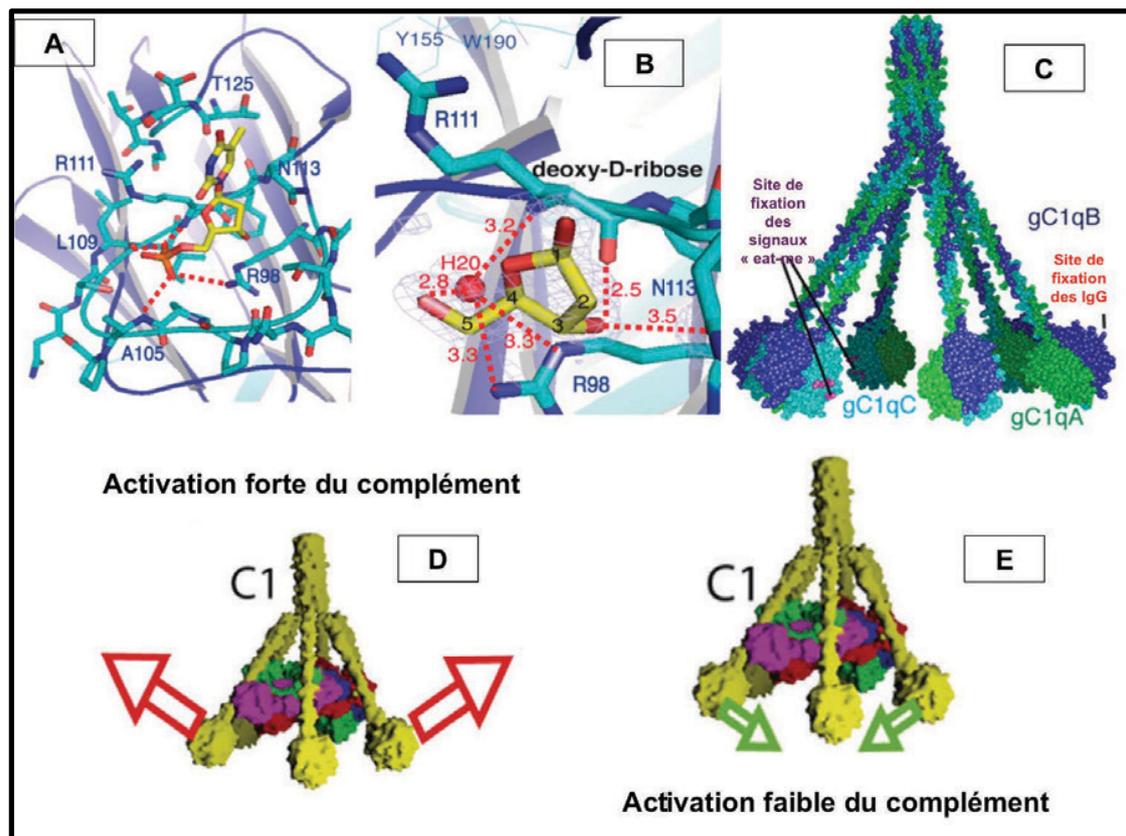


FIGURE 1.I.8 : SITE DE FIXATION DU DEOXY-D-RIBOSE, D'UN HEPARANE SULFATE ET MECANISMES D'ACTIVATION DU COMPLEXE C1 (GABORIAUD ET AL., 2011A; GARLATTI ET AL., 2010)

La résolution de la structure des GRC1q en complexe avec la phosphatidylsérine et le deoxy-D-ribose permet de proposer leur site de fixation. **(A)** L'héparane sulfate est stabilisé par les résidus R111, L109, A105, R98, N113 et T125 de la chaîne C. **(B)** Le D- déoxyribose est lié par la chaîne principale des résidus R111 et N113, ainsi que la chaîne latérale de la R98 de la chaîne C. **(C)** Les sites de fixation des signaux « eat-me » sont localisés au niveau de la chaîne C sur la face interne de la molécule C1q, alors que le site de fixation des IgG serait situé sur la face externe de la chaîne B (située à l'extérieur du cône). Ces observations ont permis de suggérer deux scénarios : la reconnaissance d'un ligand sur la face externe des GRC1q entraînerait l'ouverture de C1q et l'activation mécanique des protéases, ce qui se traduit par une activation forte du complément **(D)**, alors que l'interaction d'un ligand sur la face interne des GRC1q limiterait la possibilité d'ouverture de C1q, et donc l'activation du complément **(E)**.

II. La calréticuline, une protéine multifonctionnelle

II.1. Rôle de la calréticuline dans l'homéostasie du calcium

La calréticuline, isolée pour la première fois dans le réticulum endoplasmique (RE) d'une cellule musculaire de lapin, porte son nom du fait de sa fonction première comme protéine régulatrice du calcium. Le calcium a un rôle important dans les cellules comme une molécule universelle de la signalisation. La majorité du calcium est stocké dans le lumen du RE. Des déficiences en calcium conduisent à des perturbations du trafic entre le RE et l'appareil de Golgi, entravent le transport des protéines à travers les pores nucléaires, la transduction des signaux et la fonction chaperonne. La concentration en calcium se situe aux alentours de 2 mM dans le milieu extracellulaire, de 100 nM dans le cytoplasme, alors que dans le réticulum elle est au total de 1 mM, dont seulement 200 μ M sous sa forme libre. La calréticuline agit comme tampon du calcium à travers 43 des 82 derniers acides aminés constituant l'hélice C-terminale qui fixent 25 moles de calcium par mole de protéine avec une faible affinité ($K_D=2$ mM) et capturent 50% du calcium du RE. Elle contribue à la régulation des mécanismes SOCI (store operated calcium influx) et une fluctuation de la concentration de CRT impacte grandement l'homéostasie du calcium qui, contrairement à la fonction chaperonne, n'est pas compensée par la calnexine (Arnaudeau et al., 2002; Bastianutto et al., 1995; Mery et al., 1996; Stevens and Argon, 1999).

II.2 Contrôle qualité du repliement des protéines dans le réticulum endoplasmique : le couple calnexine-calréticuline

II.2.a. Le réticulum endoplasmique : siège du repliement des protéines

Le réticulum endoplasmique (RE) est l'organite central de la voie de sécrétion des protéines dans la cellule et à l'extérieur. Il est constitué d'une double membrane plasmique directement dérivée d'une invagination de la membrane plasmique plasmatique, ce qui fait que l'intérieur du RE est proche de l'extérieur de la cellule. Ce check-point dans la voie de sécrétion fournit la possibilité de modifications post-traductionnelles et également un contrôle qualité de haut niveau, à l'inverse d'une sécrétion directe à travers la membrane depuis le cytosol comme chez les procaryotes (Braakman and Bulleid, 2011). Il est également impliqué dans la régulation du calcium, la signalisation entre les cellules et les organelles, et la signalisation dépendante du calcium. C'est un organite unique pouvant détecter et intégrer des signaux, modifier sa propre dynamique et générer une réponse (Michalak et al., 2009).

La sécrétion de protéines fonctionnelles est essentielle à la vie chez les organismes unicellulaires et multicellulaires. Le bon fonctionnement des protéines est caractérisé par une séquence primaire en acides aminés correcte, directement issue de l'information génétique portée par l'ADN, ainsi que par l'acquisition d'une structure tridimensionnelle offrant à la protéine la possibilité d'assurer sa fonction. Au niveau de l'ADN de nombreux mécanismes de correction et de réparation maintiennent les erreurs en dessous de 10^{-8} chez les bactéries et 10^{-10} chez les eucaryotes (**Ogle and Ramakrishnan, 2005**).

Différentes études ont montré qu'une protéine seule *in vitro* dans un tube à essai possédait toute l'information nécessaire dans sa séquence primaire pour se replier correctement. La première étape du repliement consiste à « enterrer » les patches d'acides aminés non polaires (« hydrophobic collapse »), puis les interactions ioniques et la formation de pont disulfures limiteront spontanément les possibilités de conformation (**Anfinsen, 1973; Anfinsen and Haber, 1961; Anfinsen et al., 1961; Dill, 1985; White, 1961**).

Dans la cellule, 30% des protéines sont synthétisées dans le RE. Contrairement aux protéines du cytosol, elles sont traduites par des ribosomes liés à la membrane du réticulum (**Ganoza and Williams, 1969; Redman, 1969**). Pour être dirigées vers le RE, ces protéines possèdent un peptide signal. Ce signal est reconnu par les particules SRP (Particules de Reconnaissance du Signal) qui se fixent sur la séquence émergeant du ribosome, ralentissant la traduction jusqu'à ce que le ribosome se lie au canal transmembranaire composé par les deux hétéro-trimères Sec61 $\alpha\beta\gamma$: le translocon (**Görlich and Rapoport, 1993; Mitra et al., 2005; Walter and Blobel, 1980, 1982**). Une fois la reprise de la traduction engagée, le peptide signal est éliminé et la chaîne polypeptidique en cours de synthèse pénètre dans le lumen du RE où elle sera prise en charge par des protéines chaperonnes (**Gilmore et al., 1982; Meyer and Dobberstein, 1980; Walter and Blobel, 1981**), (**Hebert and Molinari, 2007**) (**Figure 1.II.1**).

Il existe différents types de protéines chaperonnes classées selon leurs modes de fonctionnement. Elles sont présentes à haute concentration dans tous les compartiments cellulaires où les phénomènes de translocation de protéines se produisent (cytosol, mitochondrie, chloroplaste, RE) et réduisent l'agrégation, les mauvais repliements et aident à la maturation. La première famille regroupe des protéines reconnaissant l'exposition de régions hydrophobes et à laquelle appartient BiP/GRP78. BiP régule l'entrée des protéines dans le RE, en bouchant le translocon et en facilitant la pénétration de la chaîne polypeptidique en cours de synthèse (**Alder et al., 2005; Hendershot, 2004; Matlack et al.,**

1999). Elle participe au repliement et à l'oligomérisation des protéines (Haas and Wabl, 1983), réduit l'agrégation des protéines non natives (Puig and Gilbert, 1994) et participe à l'homéostasie du calcium dans le RE (Lièvreumont et al., 1997). Elle joue également un rôle important dans le tri des protéines mal repliées et leur élimination dans le cytosol (Cabral et al., 2002; Kabani et al., 2003; Molinari et al., 2002; Sörgjerd et al., 2006).

La seconde famille, à laquelle appartiennent la calréticuline et la calnexine, interagit avec les protéines néosynthétisées majoritairement grâce à un site lectine qui reconnaît un motif N-glycane. Même si ces deux mécanismes de reconnaissance des protéines néosynthétisées sont différents, ils permettent de faciliter le repliement et d'améliorer la fidélité du processus de maturation (Hebert and Molinari, 2007) (Figure 1.II.1).

II.2.b. Le cycle calnexine-calréticuline

Après l'entrée de la chaîne polypeptidique en cours de synthèse dans le réticulum endoplasmique, régulée par la protéine BiP, la protéine est glycosylée. La glycosylation commence du côté cytoplasmique du réticulum endoplasmique par le transfert du groupement 2-N-acétylglucosamine décoré de 5 résidus mannose sur une ancre lipidique fixée à la membrane (dolichyl-phosphate). Sous cette forme l'ancre bascule sur la face de la membrane plasmique exposée au lumen. Ici 4 résidus mannose et 3 glucoses sont ajoutés pour obtenir le produit final $\text{Glc}_3\text{Man}_9\text{GlcNAc}_2$. Ce sucre complexe est ensuite transféré en formant une liaison amide sur un résidu asparagine de la protéine par une oligosaccharyltransférase (OST) qui reconnaît une séquence spécifique de la protéine : NXS/T (Parodi et al., 1972). Statistiquement, deux tiers des sites libres sont glycosylés, en particulier dans des zones où les structures secondaires seront modifiées (Figure 1.II.1).

Dans le lumen, la glucosidase I (GI) et la glucosidase II (GII) éliminent les deux premiers résidus glucose et génèrent le glycane monoglucosylé ($\text{Glc}_1\text{Man}_9\text{GlcNAc}_2$) reconnu spécifiquement par le domaine lectine de la CRT et de la CNX (Hammond et al., 1994; Hebert and Molinari, 2007; Kapoor et al., 2003; Spiro et al., 1996; Ware et al., 1995; Zapun et al., 1999) (Figure 1.II.1 & Figure 1.II.2 (1)).

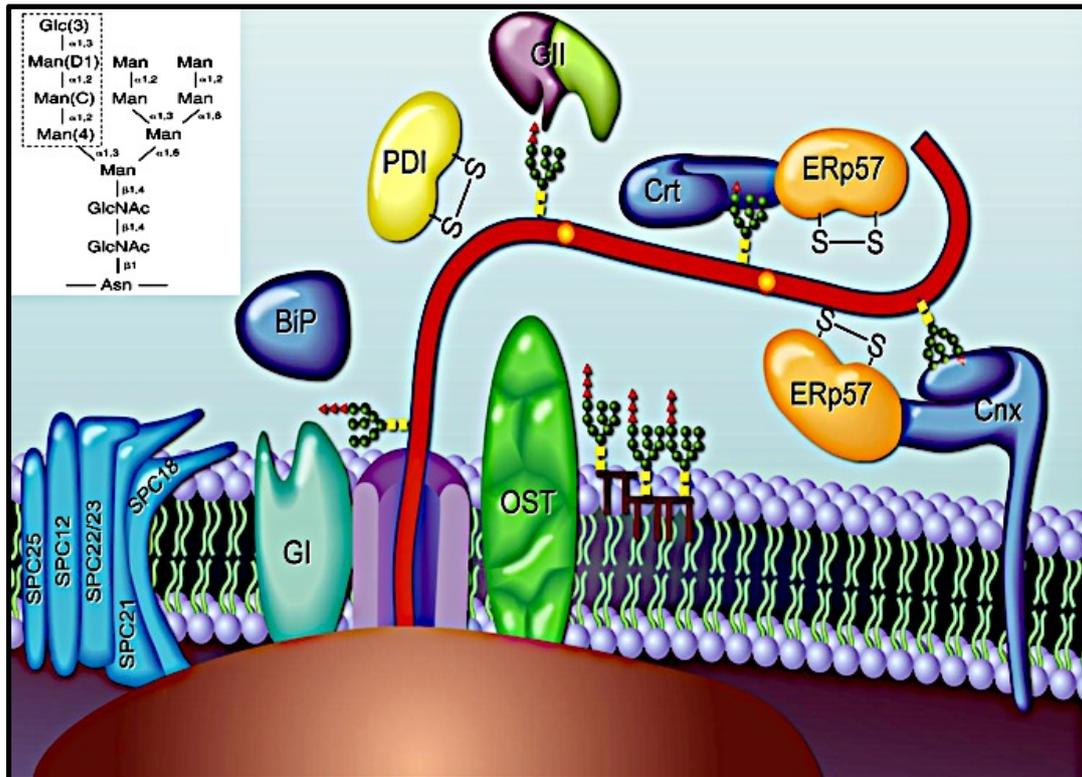


FIGURE 1.II.1 : PRISE EN CHARGE DES PROTEINES EN COURS DE SYNTHÈSE DANS LE RIBOSOME (HEBERT AND MOLINARI, 2007)

Le ribosome interagit avec le translocon (**en violet**) et la chaîne polypeptidique en cours de synthèse pénètre dans le lumen du réticulum où elle sera glycosylée par oligosaccharyltransférase (**OST**), et prise en charge par différentes chaperonnes au cours du cycle calnexine-calréticuline (**CRT-CNX**). Durant le cycle d'autres protéines interviennent, comme les glucosidases I et II (**GI et GII**) et les protéines di-sulfide isomérases (**PDI**) comme ERp57. La protéine BiP joue le rôle de bouchon à l'entrée du translocon.

La fixation aux protéines chaperonnes permet de ralentir le processus de repliement pour le rendre plus efficace (**Helenius and Aebi, 2004**) et limite les agrégations (**Hebert et al., 1996; Jackson et al., 1994; Molinari et al., 2004; Vassilakos et al., 1996**). Elle permet également la présentation de la protéine néosynthétisée aux protéines de la famille des PDI (protéines di-sulfide isomérase) à laquelle appartient ERp57, liée physiquement à la CRT et à la calnexine. Ces PDI sont responsables de la formation et de l'isomérisation des ponts disulfures nécessaires au repliement de la protéine. La topologie différente des deux lectines chaperonnes permet une large variété de substrats. La calnexine, possédant un domaine transmembranaire, fixe les glycanes à proximité de la membrane, tandis que la calréticuline, soluble, fixe les glycanes plus profondément dans le lumen du RE (**Hebert and Molinari, 2007**).

Une fois le substrat relargué par la calréticuline et la calnexine, la glucosidase II supprime le troisième et dernier résidu glucose, empêchant la protéine de se fixer sur les lectines (**Figure 1.II.2 (2), (2a), (3) et (3a)**). Si la protéine n'est pas correctement repliée, elle est prise en charge par la UDP-glucose:glycoprotein glucosyltransferase (UGGT), qui agit comme un détecteur de mauvais repliement (folding-sensor), et qui ajoute un résidu glucose pour permettre à la protéine de se fixer à nouveau au couple de lectines CNX/CRT (**Figure 1.II.2 (4a)**).

Si les protéines sont bien repliées, elles n'exposent plus de zone hydrophobe au solvant, ne sont donc ni reconnues par les chaperonnes, ni re-glycolysées. Elles seront partiellement démannosylées (ER α 1,2 mannosidase I) sur le bras B et éventuellement exportées par les voies de sécrétion. La plupart des protéines ayant subi correctement tout le processus de maturation quittent très vite le RE (**Figure 1.II.2 (4)**).

Dans le cas d'un mauvais repliement, les mannoses supplémentaires de la branche A sont enlevés (éliminant le site pouvant être glycosylé). Cette étape supplémentaire provoque l'aiguillage vers la machinerie de dégradation ERAD (Dégradation Associée au Réticulum Endoplasmique) (**Hebert and Molinari, 2007**) mais également l'impossibilité d'interagir avec les récepteurs cargos conduisant à la voie de sécrétion des protéines natives (**Figure 1.II.2 (3b) et (4b)**).

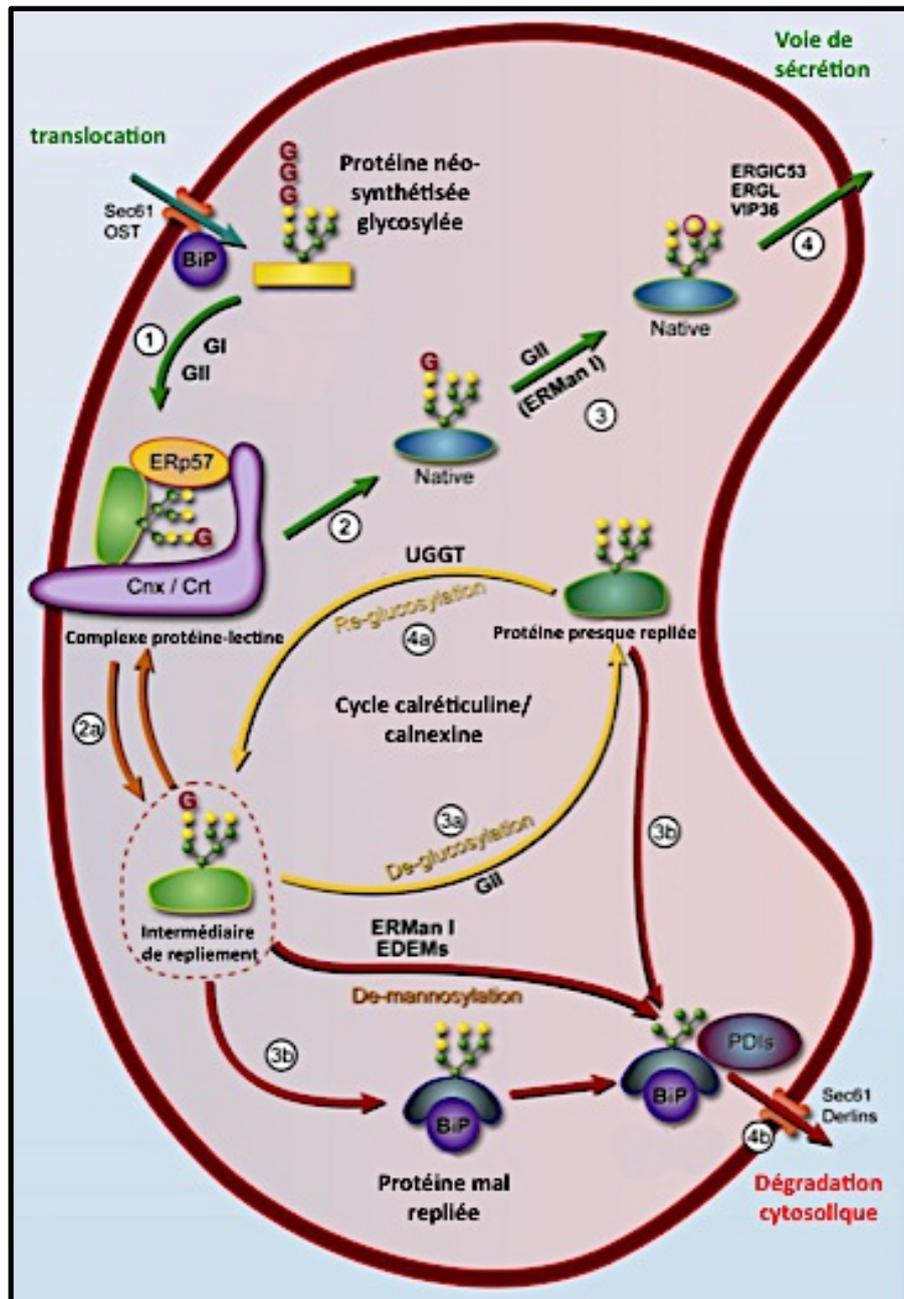


FIGURE 1.II.2 : DESCRIPTION DU CYCLE CALNEXINE-CALRETICULINE (HEBERT AND MOLINARI, 2007)

II.2.c. Structure modulaire de la calréticuline et de la calnexine

La calréticuline et la calnexine possèdent une structure modulaire similaire. Elles s'organisent autour d'un domaine globulaire N-terminal dont la structure a été obtenue par cristallographie aux rayons X. Il est constitué de 7 brins β anti-parallèles formant un feuillet concave et 6 brins β formant un feuillet convexe, le tout constituant un repliement de type « legume lectin ». Ces deux feuilletts sont maintenus par des interactions hydrophobes (Michalak et al.,

2002; Schrag et al., 2001). Le feuillet β concave des deux protéines présente un site de fixation spécifique des glycanes monoglucosylés que je décrirai dans la section suivante (**Kapoor et al., 2004; Leach et al., 2002**). Ce domaine possède un site de fixation de haute affinité du calcium, essentiel à la fonction chaperonne et à la stabilité de la protéine (le domaine est résistant à la protéolyse en présence de calcium) (**Corbett et al., 2000**), un site de fixation du zinc (**Baksh et al., 1995**), et un pont disulfure essentiel au maintien de la fonction chaperonne (**Andrin et al., 2000**) (**Kozlov et al., 2010a**). Il a été montré plus récemment que ce domaine était également impliqué dans la fixation de peptides hydrophobes, dont je décrirai les bases moléculaires dans une prochaine section (**Chouquet et al., 2011**).

Le domaine globulaire se prolonge en une structure en forme d'épingle à cheveux, nommée le domaine P (pour domaine riche en proline). La structure de ce domaine a été obtenue par cristallographie aux rayons X pour la calnexine et par RMN pour la calréticuline. Il est 25% plus long pour la calnexine et constitué de paires de répétitions (répétition A : IXDPXADXKPEDWDK / répétition B : GXWXPPXIXNPXYX), répétées 3 fois dans la calréticuline et 4 dans la calnexine (**Fliegel et al., 1989; Smith and Koch, 1989; Wada et al., 1991**). Le domaine P contient une région étendue stabilisée par 3 paires de brins β antiparallèles. Il interagit avec la protéine ERp57, plus particulièrement à travers les résidus E239, D241, E243 et W244 pour la calréticuline (**Ellgaard et al., 2001a, 2001b, 2002; Frickel et al., 2002; Martin et al., 2006; Oliver et al., 1997; Pollock et al., 2004; Schrag et al., 2001**), ainsi qu'avec la cyclophiline B (**Kozlov et al., 2010b**). Ce domaine fixe également le calcium avec une forte affinité ($K_D=1 \mu\text{M}$), mais une faible capacité (1 mol de calcium /molécule) (**Baksh and Michalak, 1991; Tjoelker et al., 1994**).

La fin du domaine P se réinsère dans le domaine globulaire qui se termine par une hélice α C-terminale comprenant de nombreuses charges négatives permettant la fixation du calcium avec une importante capacité (25 moles par molécule) mais une faible affinité ($K_D=2 \text{ mM}$). Grâce à ce domaine, la calréticuline peut fixer 50% du calcium contenu dans le lumen du RE (**Chouquet et al., 2011; Kozlov et al., 2010a; Nakamura et al., 2001**).

Contrairement à la calréticuline, cette séquence C-terminale forme une hélice transmembranaire pour la calnexine, incluant une séquence de rétention dans le RE. Ici, l'hélice fixe le calcium avec une affinité modérée. (**Michalak et al., 2009; Schrag et al., 2001; Tjoelker et al., 1994; Wada et al., 1991**) (**Figure 1.II.3**).

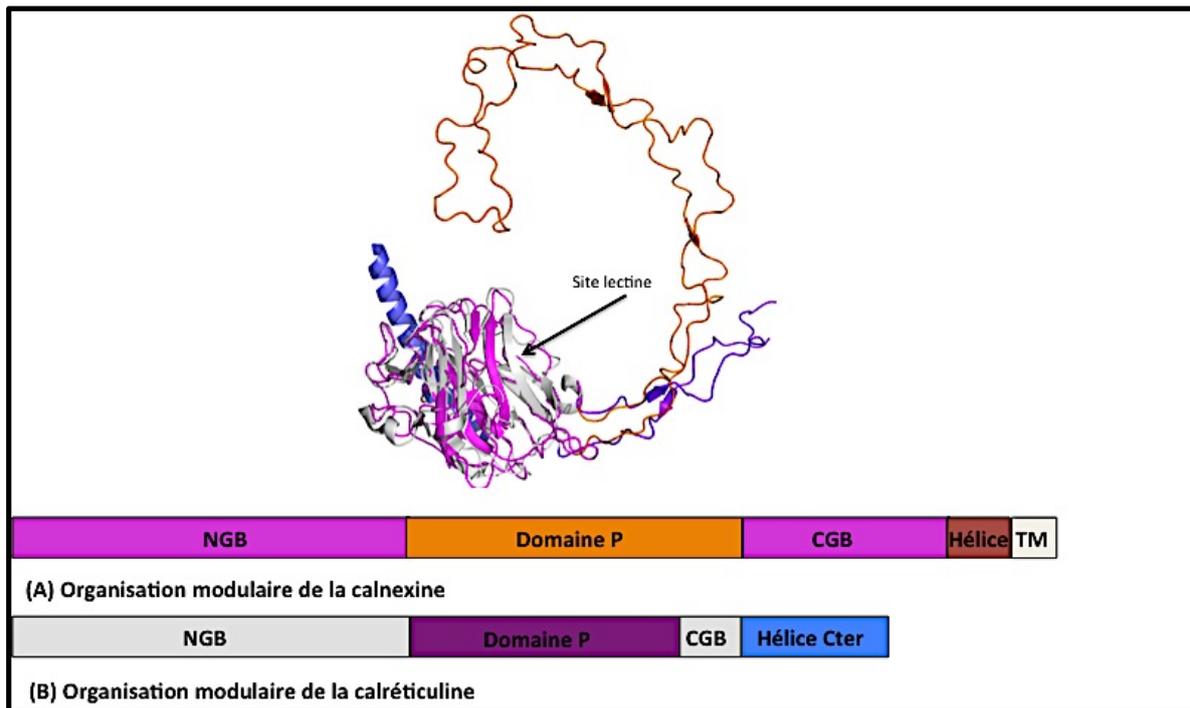


FIGURE 1.II.3 : STRUCTURE MODULAIRE DE LA CALRETICULINE ET DE LA CALNEXINE (CHOUQUET ET AL., 2011; ELLGAARD ET AL., 2002; KOZLOV ET AL., 2010A; RAGHAVAN ET AL., 2013; SCHRAG ET AL., 2001)

L'organisation modulaire de la calnexine et de la calréticuline est présentée en A et B respectivement. Le code couleur utilisé pour décrire le gène correspond à celui utilisé dans la structure protéique de la calnexine et de la calréticuline de souris.

Les protéines s'organisent autour d'un domaine globulaire (NGB) impliqué dans la fixation du calcium et du motif glycané. Le domaine P, intervenant dans la fonction chaperonne en fixant ERp57 et la cyclophiline B, est 25% plus long pour la calnexine. Il est ici tronqué dans la structure de la calréticuline. Enfin une hélice C-terminale est très importante pour la régulation de l'homéostasie du calcium dans le RE et précède une région transmembranaire pour la calnexine (TM). La taille relative des gènes et des différents domaines est respectée entre les protéines.

II.2.d. Site de fixation du glycané dans la calréticuline

L'activité chaperonne de la calréticuline et de la calnexine s'effectue principalement par la reconnaissance spécifique de substrats monoglucosylés. La résolution des structures des calréticulines humaine et de souris a permis de décrypter l'organisation moléculaire du domaine lectine (résidus 18 à 202) et plus particulièrement du site de fixation du tétrasaccharide Glc_1Man_3 (correspondant à la branche de $\text{Glc}_1\text{Man}_9\text{GlcNAc}_2$ se fixant au domaine lectine).

Le tétrasaccharide se positionne le long du canal formé par le feuillet β concave, avec une stabilisation du glucose terminal dans une cavité peu profonde formée par les chaînes latérales de la méthionine 131 et de l'isoleucine 147. De nombreuses liaisons hydrogène stabilisent le

glucose dans cette cavité : l'oxygène O2 établit des liaisons avec la chaîne latérale de la lysine 111, le carbonyle du squelette de l'asparagine 154, tandis que les oxygènes O4 et O6 sont stabilisés par les chaînes latérales de la tyrosine 109 et de l'aspartate 135, respectivement **(Kozlov et al., 2010a) (Figure 1.II.4)**.

La structure cristallographique permet également de comprendre la spécificité pour un glucose terminal, et non pas un résidu mannose ou trois glucoses. En effet le glucose et le mannose sont des épimères au niveau du carbone C2, et, alors que l'oxygène équatorial O2 du glucose s'insère parfaitement dans le sillon formé par les chaînes latérales, le mannose aurait entraîné un encombrement stérique avec la méthionine 131, ainsi que la perte de deux liaisons hydrogène (avec le carbonyle de la glycine 124 et la chaîne latérale de la lysine 111) **(Kapoor et al., 2004; Kozlov et al., 2010a; Thomson and Williams, 2005) (Figure 1.II.4)**.

La stabilisation du résidu mannose, en position D1, s'effectue quant à elle majoritairement à travers des liaisons hydrogène impliquant les oxygènes O4 et O6, avec la chaîne latérale de la tyrosine 109, les chaînes latérale et principale de l'aspartate 317. De leur côté, les atomes O4 et O6 du résidu mannose en position C sont stabilisés par la chaîne latérale de l'aspartate 135, la chaîne latérale de la tyrosine 109 et le groupement amide du tryptophane 319. Les chaînes latérales de l'aspartate 135 et de la tyrosine 109 interviennent également dans la stabilisation de l'oxygène O3 par l'intermédiaire de deux molécules d'eau **(Kozlov et al., 2010a) (Figure 1.II.4)**.

Ces différentes observations concordent avec la diminution de la fixation du carbohydrate observée lors de la mutation des résidus W319 et M131 en alanine, alors que la mutation de D125 ou D160 ne montre aucun effet. L'importance du pont disulfure entre les résidus cystéine 105 et 137 a également un rôle essentiel pour la fonction chaperonne et la fixation du carbohydrate. La structure démontre que cet effet est dû en partie au contact établi entre la liaison C5-C6 du mannose C et la liaison C1-O1 du mannose 4 qui enveloppent le pont disulfure **(Gopalakrishnapai et al., 2006; Kapoor et al., 2004; Kozlov et al., 2010a; Martin et al., 2006; Thomson and Williams, 2005)**.

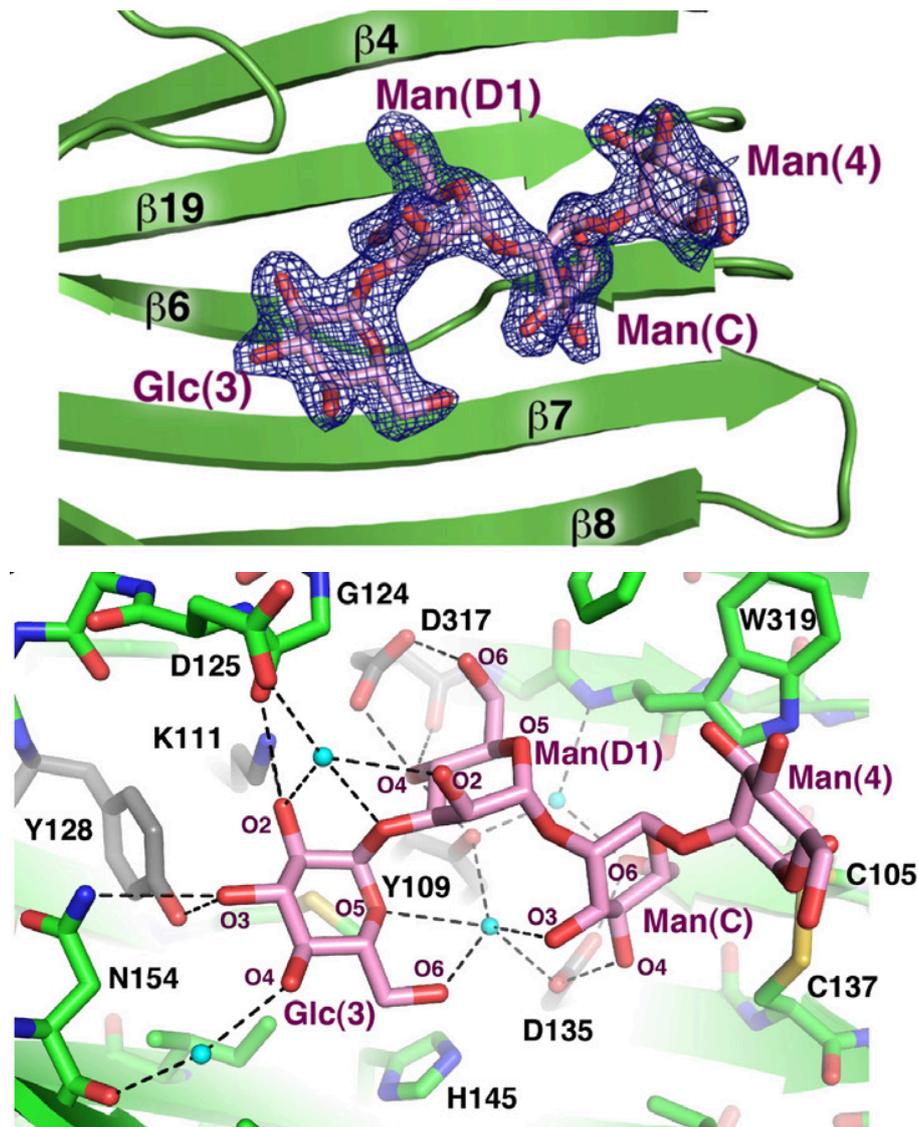


FIGURE 1.II.4 : CARTOGRAPHIE STRUCTURALE DU SITE DE RECONNAISSANCE DU CARBOHYDRATE DANS LE SITE LECTINE DE LA CALRETICULINE (KOZLOV ET AL., 2010A)

(a) La carte OMIT calculée en absence de ligand modélisé, montre la fixation du substrat Glc_1Man_3 dans le feuillet β concave. (b) Le carbohydrate est stabilisé par des liaisons hydrogène principalement avec les résidus Y109, D125, Y128, D135, N154 et D317 et trois molécules d'eau.

Le domaine lectine est extrêmement conservé entre la calréticulines et la calnexine (33% d'identité et une déviation de 1,7 Å sur les positions de 195 atomes CA). La similarité est plus marquée au niveau des feuillets β . Des différences structurales sont trouvées dans la position des différentes boucles reliant les brins β entre eux, avec également l'absence de 3 hélices dans la calnexine, se situant le long du feuillet β convexe dans la calréticuline (A32-

R36, L190-D199, E336-D362) (Chouquet et al., 2011; Kozlov et al., 2010a; Schrag et al., 2001).

Les résidus impliqués dans la fixation du tétrasaccharide sont extrêmement conservés et adoptent des conformations similaires. En effet le glucose dans la structure de la calnexine est stabilisé par des liaisons hydrogène avec les résidus Y165, K167, Y186, E217 et E426, les équivalents dans la calréticuline de Y109, K111, Y128, N154 et D317 respectivement (Figure 1.II.4 et 1.II.5). Dans les calréticulines, un ion calcium est coordonné par D326, Q26, K62, K64 et deux molécules d'eau. Il n'est pas visible dans la structure de la calnexine à cause de la qualité insuffisante des données, mais les résidus D437, S75, D118 et K119 aux positions équivalentes sont bien présents (Chouquet et al., 2011; Kozlov et al., 2010a; Leach and Williams, 2004; Schrag et al., 2001) (Figure 1.II.6).

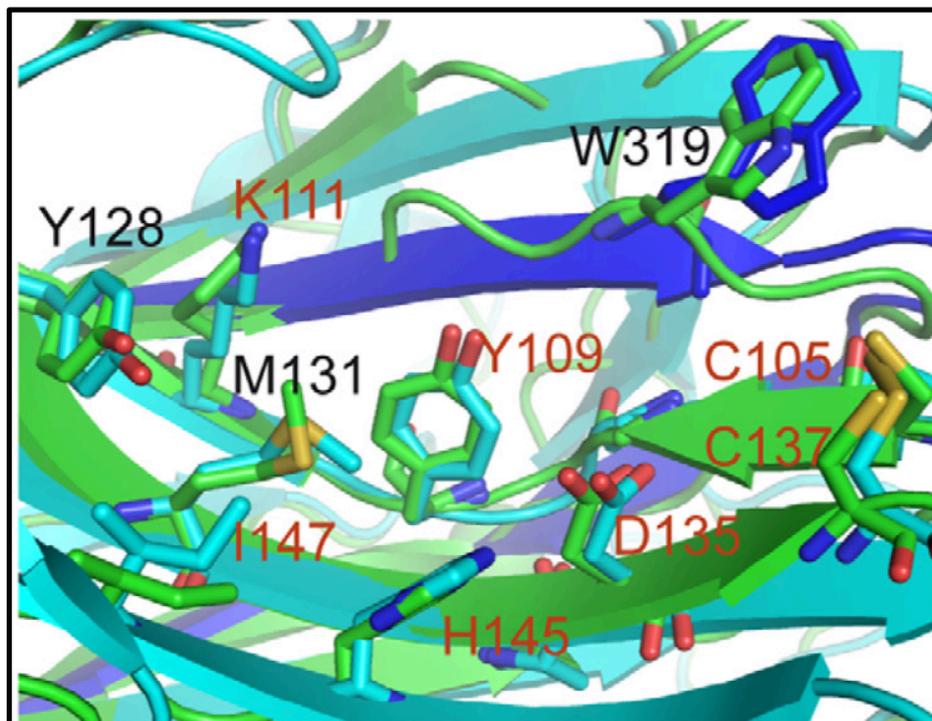


FIGURE 1.II.5 : SUPERPOSITION DU SITE LECTINE DE LA CALRETICULINE HUMAINE ET DE LA CALNEXINE DE CHIEN (CHOUQUET ET AL., 2011; SCHRAG ET AL., 2001)

La superposition des structures obtenues par cristallographie aux rayons X de la calréticuline humaine (cyan) et de la calnexine (vert) montre une importante conservation du site lectine, impliqué notamment dans la fixation du motif carbohydate (seuls les acides aminés strictement conservés sont montrés).

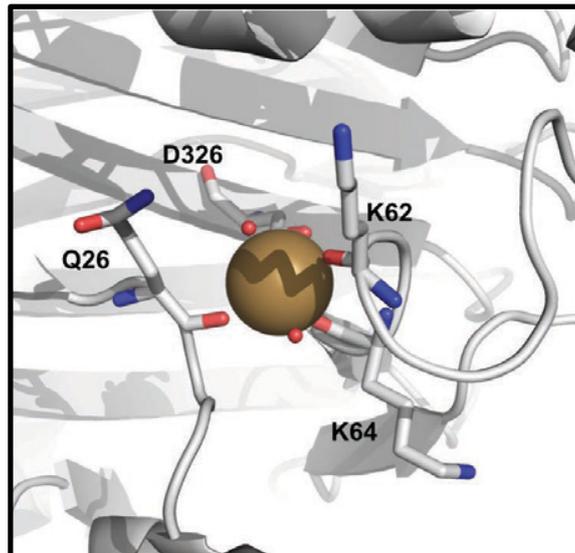


FIGURE 1.II.6 : SITE DE FIXATION DU Ca^{2+} DANS LES CALRETICULINES DE MAMMIFERE (CHOUQUET ET AL., 2011)

L'atome de calcium est stabilisé par des liaisons hydrogène avec les chaînes principales des résidus D326, Q26, K62, K64 et deux molécules d'eau.

II.2.e Fixation peptidique indépendante du glycane dans la calréticuline

De manière surprenante, l'absence de la calréticuline affecte également le repliement de protéines non glucosylées comme le montre le rôle de la CRT sur la maturation du complexe majeur d'histocompatibilité (CMH) de classe I (**Michalak et al., 2009; Rutkevich and Williams, 2011**). Les CMH sont des acteurs majeurs de l'immunosurveillance de l'organisme. Ils présentent une étiquette peptidique, provenant de peptides de pathogènes intracellulaires ou de cellules transformées de l'organisme, aux cellules immunitaires (comme les LT CD8⁺). Les CMH sont composés d'une chaîne lourde α polymorphique et d'une chaîne β_2 -microglobuline. Sous cette forme, ils sont très instables. Ils seront stabilisés et spécialisés par l'ajout du peptide au sein d'un domaine structural unique à chaque CMH : le sillon de fixation du peptide. Le chargement du peptide s'effectue par un complexe protéique incorporé à la membrane du réticulum endoplasmique, le PLC (peptide-loading complexe) constitué de la tapasine, des protéines TAP (Protéines de Transport des Antigènes), de ERp57 et de la CRT. L'absence de la calréticuline diminue la stabilité et la maturation du CMH en empêchant le chargement du peptide. Cela se caractérise également par un export plus rapide du CMH à la surface cellulaire, montrant que, en plus d'intervenir dans le complexe PLC, la calréticuline intervient également dans la rétention du complexe non mature (**Elliott and Williams, 2005; Hammer and Shastri, 2007**).

Ces différentes observations sont cohérentes avec le rôle de protéine chaperonne porté par la calréticuline. Cependant certaines études montrent que l'inactivation de la fonction lectine de la calréticuline par mutation ne perturbe pas l'assemblage des CMH et sa capacité à charger le peptide (**Ireland et al., 2008; Leach and Williams, 2004**). Et même si d'autres travaux mettent en lumière une perte de 30 à 50% de l'activité de la calréticuline dans la maturation des CMH (**Del Cid et al., 2010**), ces constatations semblent confirmer un concept proposé dès 1995, d'une dualité de fixation faisant intervenir un site de fixation peptidique agissant de concert avec le site de fixation du glycane (**Ihara et al., 1999; Ware et al., 1995**). Depuis, cette hypothèse a été largement appuyée par la mise en évidence du rôle de la calréticuline pour interagir avec de nombreuses protéines non-glycosylées et diminuer leur agrégation (**Brockmeier and Williams, 2006; Del Cid et al., 2010; Duus et al., 2009; Ihara et al., 1999; Jeffery et al., 2011; Martin et al., 2006; Pocanschi et al., 2011; Saito et al., 1999**). De même, la fonction chaperonne est inhibée en mutant le résidu H153 alors que celui-ci se trouve à 20 Å du site lectine (**Guo et al., 2003**).

Actuellement, un des enjeux majeurs est de déterminer précisément le site d'interaction des peptides et de définir dans quel contexte ce mode de fixation prend le dessus sur la fonction lectine (ou vice versa). Pour répondre à cette seconde question, de nombreux travaux ont été réalisés sur la stabilité de différents mutants de calréticuline et sur les conditions dans lesquelles l'interaction mettant en jeu des peptides serait privilégiée *in vivo*.

Il a été montré qu'une déstabilisation de la calréticuline avait un impact sur sa fonction chaperonne (**Del Cid et al., 2010; Ihara et al., 1999; Ireland et al., 2008; Leach et al., 2002; Martin et al., 2006; Rizvi et al., 2004; Thammavongsa et al., 2005**). La fixation indépendante du glycane serait prépondérante dans des conditions de stress du RE, comme par exemple une diminution de la concentration en calcium ou un choc thermique (**Rizvi et al., 2004**). Ces observations sont directement liées à un changement de conformation de la protéine, correspondant à une déstabilisation et l'exposition de résidus à caractère hydrophobe (**Jeffery et al., 2011**).

Comme pour la calnexine, l'utilisation de fragments de calréticuline et de peptides hydrophobes dans des expériences de compétition a montré que le site de fixation indépendant du glycane se trouvait également dans le domaine globulaire, car aucune compétition n'est observée avec les peptides après ajout du substrat Glc₁Man₃ (**Pocanschi et al., 2011**). Cependant, la fonction chaperonne associée impliquerait également le domaine P,

contrairement au mécanisme dépendant du glycane (Del Cid et al., 2010; Jeffery et al., 2011; Schrag et al., 2001).

L'analyse des structures des calréticulines humaine et murine a conduit à suggérer un site potentiel de fixation des peptides sur le bord du site lectine. Ce site impliquerait les résidus F74, Y109, C105, C137, D135 et W319 comme illustré sur la **figure 1.II.7**. Dans les différents empilements cristallins obtenus, ce site interagit avec l'extrémité N-terminale d'une molécule voisine. Cette extrémité flexible correspond à un ajout exogène introduit pour la production et purification des protéines recombinantes, et donc peut être assimilée à un peptide non replié. La **figure 1.II.7** montre l'exemple d'interaction observé dans la 1^{ère} forme cristalline de la CRT humaine. Des contacts additionnels hydrophobes ont été observés dans d'autres empilements cristallins avec F46, M131, D317, G106 et G124 (Chouquet et al., 2011; Kozlov et al., 2010a) (**Figure 1.II.7**).

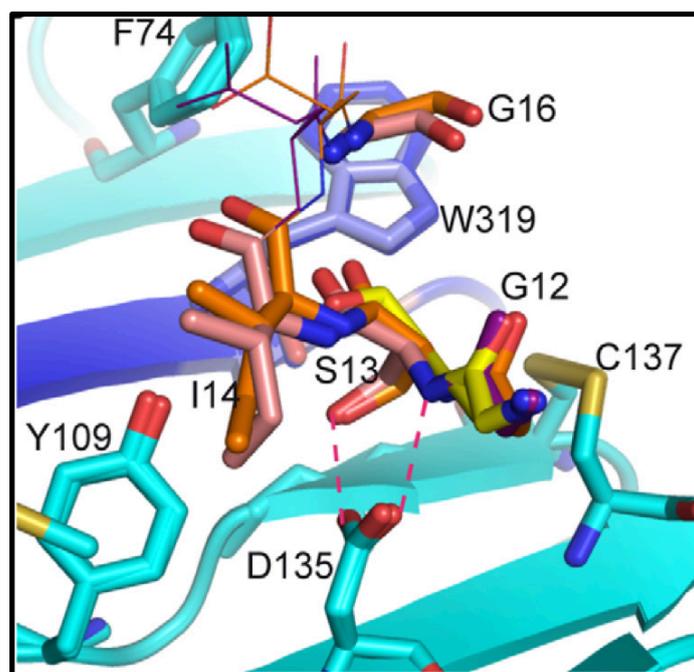


FIGURE 1.II.7: SITE POTENTIEL D'INTERACTION DU PEPTIDE SUR LE BORD DU DOMAINE LECTINE DE LA CALRETICULINE (CHOUQUET ET AL., 2011).

La stabilisation de l'extrémité N-terminale ajoutée artificiellement pour la purification de la calréticuline humaine par le bord du domaine lectine dans l'empilement cristallin à travers des interactions majoritairement hydrophobes met en évidence un nouveau site potentiel d'interaction avec des peptides. Ce site présente le même type d'interactions dans la structure de la calréticuline de souris (non présenté) et implique les résidus F74, Y109, C105, C137, D135 et W319.

Cette hypothèse est confortée par le fait que les mutations de résidus uniquement impliqués dans la fixation du glucose terminal (Y109, K111, Y128, D317) sont sans effet sur la fonction chaperonne indépendante du glycane (Ireland et al., 2008; Thomson and Williams, 2005), au contraire des mutations des deux cystéines 105 et 137 ainsi que du Trp319 (résidus inclus dans le site proposé). De plus, des expériences récentes d'atténuation de fluorescence situent clairement le site de fixation des peptides à proximité du site de fixation du glycane (Wijeyesakere et al., 2013).

Aujourd'hui, même si les preuves semblent de plus en plus localiser la fonction chaperonne indépendante du glycane sur le bord du site lectine, certaines ambiguïtés doivent être encore levées. En effet, le substrat Glc_1Man_3 n'inhibe pas la fixation de peptides hydrophobes (suggérant deux sites distincts) mais inhibe la fixation de protéines glycosylées aussi bien que non glycosylées (suggérant un site unique). Il se pourrait alors que les deux modes de fixation puissent coexister dans le cadre de petits peptides. Pour les protéines, il semblerait que le domaine P se ferme, permettant de maintenir le substrat contre le domaine lectine et semble dans ce cas jouer un rôle non négligeable dans l'interaction (Figure 1.II.8) (Wijeyesakere et al., 2013).

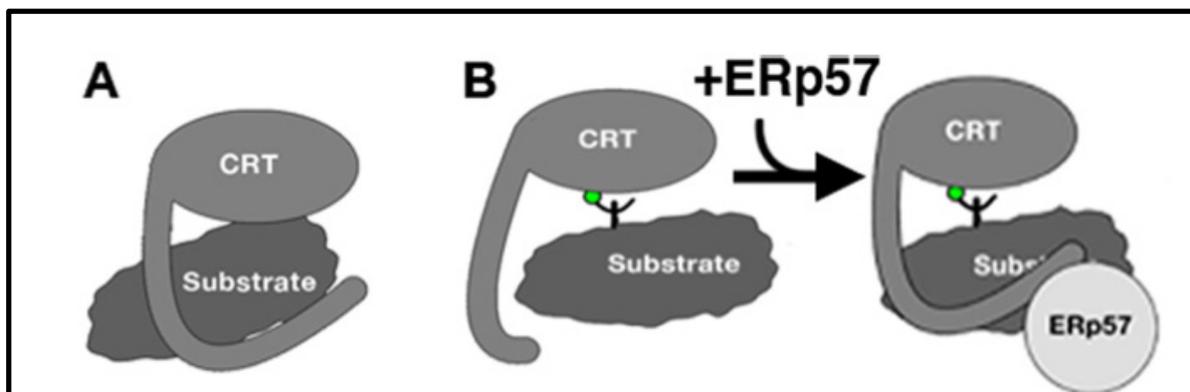


FIGURE 1.II.8 : MODELES DE FIXATION DES PROTEINES NEOSYNTHESEES AVEC LA CALRETICULINE (WIJEYSAKERE ET AL., 2013)

(A) Le mécanisme de fixation indépendant du glycane implique le domaine P qui se referme sur le substrat. (B) Au contraire, le mécanisme de reconnaissance du tétrasaccharide est indépendant du domaine P et le bloque en position ouverte. L'association de ERp57 avec le domaine P provoque sa fermeture, permettant de mettre la disulfure isomérase au contact du substrat.

II.3 Fonctions extracellulaires de la calréticuline

En plus de ses fonctions dans le réticulum endoplasmique comme protéine chaperonne et régulateur de la concentration du calcium, la calréticuline possède également de nombreuses fonctions lorsqu'elle est exportée à la surface cellulaire ou sécrétée.

II.3.a. Effets anti-angiogéniques

Sous sa forme sécrétée, la calréticuline est capable d'interagir avec les laminines et la thrombospondine. Ces protéines sont des constituants de la matrice extracellulaire et participent aux jonctions intercellulaires. La calréticuline, et plus particulièrement la région N-terminale de son domaine globulaire, peut former un complexe avec les laminines, ce qui bloque l'interaction des cellules endothéliales entre elles et réduit leur capacité à s'attacher à la membrane extracellulaire et donc de proliférer (pour former par exemple des vaisseaux sanguins) (Aguilar-Guzmán et al., 2014; Li et al., 2007; Pike et al., 1998, 1999; Ramírez-Toloza et al., 2014; Yao et al., 2002). Cet effet anti-angiogénique est très intéressant. En effet, un apport sanguin via l'angiogénèse est vital pour les cellules tumorales. La CRT semble être efficace dans la réduction des tumeurs (Cai et al., 2008; Jazowiecka-Rakus et al., 2006; Ma et al., 2007; Xiao et al., 2002), même si cet effet semble dépendant du type tissulaire puisque dans d'autres cas elle semble induire la croissance tumorale (Liu et al., 2005).

II.3.b. Implication dans le phénomène de cicatrisation

Le processus de cicatrisation se déroule en trois phases : la première est la formation de caillot de fibronectine et une agrégation des plaquettes, avec une migration de cellules pro-inflammatoires qui phagocytent les bactéries et les cellules mortes et libèrent des facteurs de croissance et angiogéniques dans la blessure. Ensuite les cellules épithéliales prolifèrent pour recouvrir la blessure et enfin des protéases comme la collagénase agissent pour remodeler la matrice extracellulaire au niveau de la blessure (Michalak et al., 2009). La calréticuline est sur-exprimée temporairement pendant la cicatrisation, mais nous ne savons pas si son action est intra ou extracellulaire. Néanmoins, un traitement de la blessure avec la calréticuline, provoque un recouvrement de la blessure par les cellules épithéliales plus rapide et avec une abondance de tissus accrue (Gold et al., 2006; Nanney et al., 2008). Cette action s'effectue à travers l'induction de la prolifération et la mobilité des kératinocytes et des fibroblastes *in vitro* et *in vivo*. La CRT stimule la prolifération des cellules endothéliales microvasculaires et donc a un effet positif dans ce cas précis sur l'angiogénèse (Nanney et al., 2008). La CRT

attire également les monocytes et macrophages pour éliminer les débris cellulaires et les bactéries.

II.3.c. Interaction avec des protéines non glycosylées : une piste pour la fonction chaperonne indépendante du glycane ?

Pour assurer ses différentes fonctions extracellulaires, notamment à travers ses interactions avec la matrice extracellulaire, la calréticuline interagit spécifiquement avec des protéines non glycosylées, comme GABARAP (GABA(A) Receptor-Associated Protein) ou la thrombospondine. Les sites d'interaction proposés pour ces deux protéines sont localisés aux résidus 35-53 et 195-205, respectivement. (Goicoechea et al., 2002; Thielmann et al., 2009) (Figure 1.II.9).

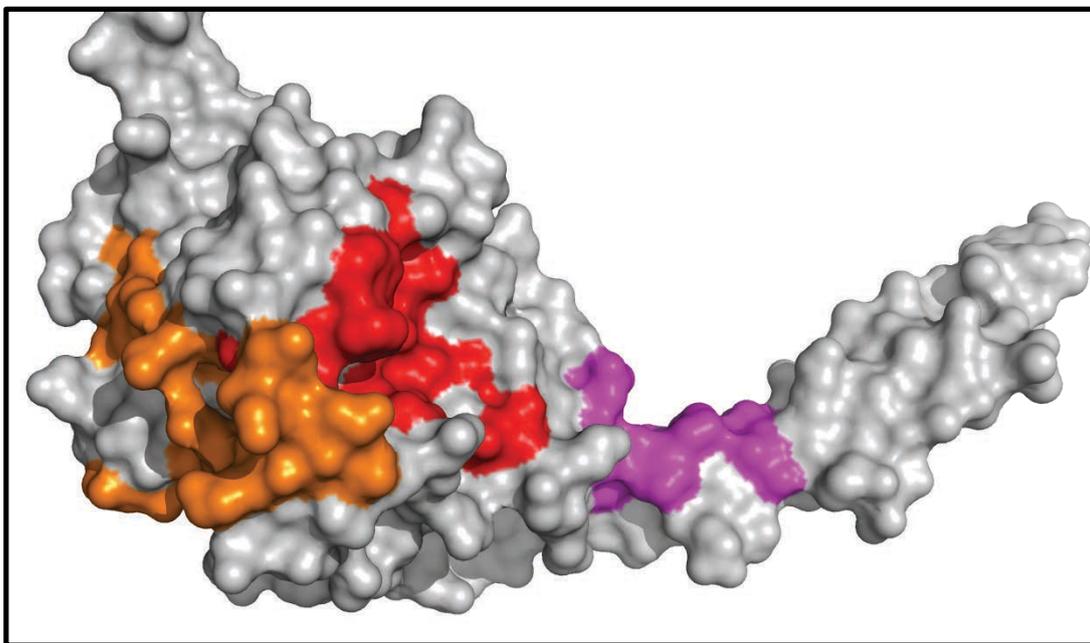


FIGURE 1.II.9: REGIONS POTENTIELLES D'INTERACTIONS PEPTIDES-DEPENDANTES

Les régions proposées comme participant à l'interaction avec la protéine GABARAP et la thrombospondine sont représentées en orange et en violet, respectivement, et sont distinctes du site lectine (en rouge).

Il a été très vite suggéré que ces sites pourraient être également impliqués dans la fonction chaperonne indépendante du glycane mais les expériences de compétition avec des peptides correspondant aux séquences de GABARAP et de la thrombospondine n'inhibent pas la capacité de la calréticuline à prévenir l'agrégation de substrats non glycosylés. Il est donc important de discriminer clairement les interactions spécifiques de protéines impliquant une

fonction extracellulaire de la calréticuline de celles moins spécifiques permettant de fixer des peptides dans le cadre de la fonction chaperonne.

II.3.d. CRT et efferocytose

Comme C1q, la calréticuline possède de nombreuses facettes fonctionnelles en lien avec le processus d'efferocytose. Tout d'abord, elle peut participer au déclenchement de l'apoptose lorsque celle-ci est induite par un stress du réticulum lié à une dérégulation des flux de calcium, puisque la CNX et la CRT s'accumulent à l'interface du contact mitochondrie-RE et participent à la transmission du calcium vers la mitochondrie (**Hayashi et al., 2009**).

Ensuite, la CRT présente à la surface de certains types cellulaires, peut être exposée plus fortement à la surface des cellules apoptotiques (**Gardai et al., 2005**), où elle joue le rôle de signal « eat-me », ce qui signifie qu'elle améliore leur élimination par phagocytose. (**Gardai et al., 2005; Obeid et al., 2007; Peters and Raghavan, 2011; Tesniere et al., 2010**). L'exposition de la calréticuline est très intéressante dans le contexte de la mort cellulaire immunogène. En effet, l'induction de l'apoptose des cellules cancéreuses avec des UV ou des molécules chimiques (comme celle de la famille de l'anthracycline), provoque la translocation de la CRT à leur surface et favorise leur élimination par les cellules dendritiques notamment. Ces observations ouvrent des perspectives intéressantes, puisqu'induire l'exposition précoce de la calréticuline à la surface des cellules cancéreuses favorise leur élimination, tandis que l'injection de ces mêmes cellules peut-être utilisée comme une vaccination anti-tumorale (**Obeid et al., 2007**). Cet export à la surface nécessite plusieurs effecteurs, comme l'activation de la kinase PERK, le clivage de la caspase 8 ou encore l'activation de Bax/Bak (**Ng et al., 1997**). La CRT n'étant pas une protéine membranaire, sa présence en surface implique des interactions moléculaires avec d'autres partenaires, comme ERp57 ou la phosphatidylsérine (PS), avec laquelle elle interagit de manière dépendante du calcium (**Gardai et al., 2005; Païdassi et al., 2011; Tarr et al., 2010**). Il faut noter cependant que l'exposition de la CRT en surface des cellules apoptotiques se produit environ 1 h avant l'exposition de la PS, un signal « eat me » emblématique (**Obeid et al., 2007**).

A la surface des macrophages, monocytes ou fibroblastes, le récepteur CD91 (LRP1) participe à la reconnaissance de la CRT, ce qui active la phagocytose (**Gardai et al., 2005**) (**Orr et al., 2002, 2003**) (**Chao et al., 2010; Gardai et al., 2003; Vandivier et al., 2002**).

Du côté du phagocyte, la présence de la CRT dans le RE est requise pour former la cupule phagocytaire dépendante de l'excroissance de filaments d'actine (**Müller-Taubenberger et**

al., 2001). La CRT semble également faciliter la phagocytose lorsqu'elle est en complexe avec CD91 à la surface des macrophages (Ogden et al., 2001; Sim et al., 1998). Nous avons vu dans le chapitre précédent que la protéine C1q, acteur majeur extracellulaire de l'immunité innée, était capable de faciliter ce processus d'efferocytose. Il a été proposé en particulier que C1q pouvait jouer le rôle de molécule pontante en interagissant avec la CRT à la surface des cellules apoptotiques et à la surface des phagocytes.

III. Stratégies d'évasion : C1q et les calréticulines de parasite

III.1. Inhibition et subversion du système du complément

La lutte entre les organismes pathogènes comme les bactéries, les virus et les parasites et le système immunitaire est une course à l'armement visant à déstabiliser l'adversaire. Certains organismes comme les bactéries jouent sur leur vitesse de multiplication afin de prendre de vitesse le système immunitaire. Dans ce cas de figure, pouvoir ralentir la cascade du complément devient un atout majeur et les différentes étapes de régulation constituent une faille importante exploitée par les organismes pathogènes. Par exemple *Staphylococcus aureus* possède une protéase capable de cliver C3 et de l'inhiber, il possède également des protéines de surface pouvant lier et cliver les Ig, les empêchant d'interagir avec C1q ou leurs récepteurs à la surface des macrophages, tandis que d'autres organismes ne possédant pas de protéase endogène utiliseront celles de l'hôte en se fixant par exemple au plasminogène pour l'activer, comme *Streptococcus pneumoniae* (Terrasse et al., 2012). Outre les étapes de reconnaissance initiale *S. aureus* peut également interférer avec l'assemblage des C3 et C5 convertases, grâce à ses protéines comme Sbi ou SSL7 formant des complexes avec C3 et C5, conduisant à l'inactivation de C3b par le facteur I et à l'inhibition de l'hydrolyse de C5. De plus, la majorité des bactéries possèdent un site de fixation pour le facteur H et peuvent recruter C4BP et le facteur CFHL-1 (Merle et al., 2015b). Ces convertases sont d'ailleurs la cible de nombreuses bactéries, puisque *Streptococcus pyogenes* inhibe la C5 convertase en utilisant sa GAPDH qui fixe C5a et ScPA l'hydrolyse. Enfin d'autres organismes agiront à un stade plus tardif en interférant avec la mise en place du complexe d'attaque membranaire (MAC), par exemple *Borrelia burgdorferi* possède deux protéines se fixant à C7 : une protéine CD59-like et CRASP-1 (Zipfel et al., 2007, 2013).

D'autres organismes, comme les virus et les pathogènes, en plus de pouvoir inhiber le système du complément, ont développé des stratégies très subtiles au cours de l'évolution pour utiliser notre système immunitaire à leur avantage. Par exemple le VIH-1 utilise la protéine C3 comme une tenue de camouflage ; en effet les dépôts de C3b et C5b à sa surface lui permettent d'interagir avec les récepteurs de l'hôte comme CR3, CR1, CR2 et CR4, mais sans être pour autant inquiété puisqu'il s'approprie des régulateurs du complément comme CD59 ou CD55 lui permettant de dégrader progressivement C3b en fragments inactifs, jusqu'à C3d. De plus il a été montré que l'interaction de C1q avec la protéine gp41 favorisait la pénétration du virus, de même que la présence de certains anticorps à sa surface. Le virus d'Epstein-Barr utilise également le récepteur CR2 pour pénétrer dans les LB (**Tanner et al., 1987; Yu et al., 2010**).

III.2. Reconnaissance des calréticulines de parasite par C1q

De nombreux organismes pathogènes utilisent la CRT lors de l'infection. Par exemple la calréticuline de *Haemonchus contortus* interagit avec C1q et la CRP de l'hôte, alors que l'ectoparasite *Amblyomma americanum* sécrète la CRT pendant qu'il se nourrit, utilisant son effet anticoagulant, ou *Schistosoma* qui l'utilise pour traverser la peau (**Dharampaul et al., 1993; Khalife et al., 1994; Naresha et al., 2009**). Ces organismes *comme les nombreux types cellulaires chez les mammifères* exposent la CRT à leur surface où elle présente souvent les mêmes fonctions comme la possibilité d'interagir avec des protéines de la matrice extracellulaire pour permettre l'adhésion des parasites aux cellules hôtes. Durant ma thèse deux parasites nous ont particulièrement intéressé : *Trypanosoma cruzi* et *Entamoeba histolytica*. En plus de posséder, comme de nombreux organismes, des éléments permettant d'inhiber le système du complément (*T. cruzi* possède par exemple une protéine T-DAF, homologue à DAF, et *E. histolytica* une adhésine spécifique du galactose lui permettant d'interférer avec la mise en place du MAC), ils semblent avoir développé une stratégie de mimétisme cellulaire (**Osorio et al., 2012; Zipfel et al., 2013**).

Ces deux organismes exposent à leur surface la calréticuline, qui a la particularité de jouer le rôle de récepteur à C1q. En plus d'inhiber la voie classique du complément pour assurer la survie des parasites, ces interactions, que je décrirai plus en profondeur dans la suite de l'introduction, permettent à *T. cruzi* de pénétrer dans les macrophages suivant un mécanisme semblable à ce que j'ai pu décrire pour les cellules apoptotiques, alors qu'elles permettent à *E. histolytica* de phagocyter les cellules de l'hôte.

IV. La calréticuline de *Trypanosoma cruzi*

IV.1. *Trypanosoma cruzi*, agent infectieux de la maladie de Chagas

La maladie de Chagas, maladie tropicale découverte en 1907, est causée par le parasite *Trypanosoma cruzi*. Elle a longtemps été considérée comme une maladie rare cantonnée aux zones rurales d'Amérique latine. Cependant les flux migratoires vers l'Amérique du nord, l'Europe et l'Asie, ainsi que le tourisme ont considérablement modifié la géographie de l'épidémie qui atteint les régions urbaines (il y aurait par exemple 300 000 personnes infectées chroniquement aux Etats-Unis). Le développement de la maladie dans des zones non endémiques pose un problème majeur de santé publique, étant donné que la plupart des individus atteints sont asymptomatiques et restent des sources de transmissions potentielles par transfusion sanguine, greffes d'organes ou via le placenta de mère à enfant (**Machado et al., 2012**).

Cette maladie progresse selon différentes phases :

- Des fièvres, vomissements, nausées, diarrhées, éruptions cutanées peuvent éventuellement apparaître au cours d'une phase aigüe qui présente des signes physiques caractéristiques: une lésion inflammatoire au niveau de la piqûre (un chagoma), un œdème périorbital et une conjonctivite (nommés signe de Romana). Chez la majorité des patients, la réponse immunitaire permet d'éliminer ces symptômes après 2 à 4 mois. Durant cette phase, la forme infectieuse trypomastigote du parasite est retrouvée dans le sang ainsi que dans le liquide cébrospinal; les cas mortels sont très rares (<2%) et concernent essentiellement les enfants. La présence du parasite peut engendrer des pathologies liées au système nerveux central (méningo-encéphalites), tandis que l'invasion des cellules gliales provoque fréquemment des dommages aux neurones, et peut également aboutir à la dilatation des organes du tractus intestinal (estomac, colon, œsophage, intestin).
- A la fin de cette période, les individus infectés entrent dans un état intermédiaire ambigu pouvant durer de quelques mois à la fin de la vie. Celui-ci se traduit par la présence du parasite dans le sang et l'absence de symptômes cliniques.
- Enfin, 30 à 40% des individus vont développer la phase chronique de la maladie. La présence du parasite provoque une inflammation massive des

tissus, notamment dans les myocytes, pouvant conduire à une insuffisance et des arrêts cardiaques (Machado et al., 2012).

IV.2. Trypanosoma cruzi

IV.2.a. Cycle de vie

Trypanosoma cruzi est un parasite obligatoire isolé pour la première fois en 1907 au Brésil par Carlos Chagas à partir d'un singe *Saimiri sciureus*. Il est capable d'infecter un nombre considérable d'espèces. Sa multiplication et sa survie s'effectuent à travers un cycle de vie très complexe décrit dans la **figure 1.IV.1** nécessitant un hôte invertébré (le vecteur) et un hôte vertébré. Les vecteurs invertébrés de *T. cruzi* sont principalement des insectes hématophages comme le triatome.

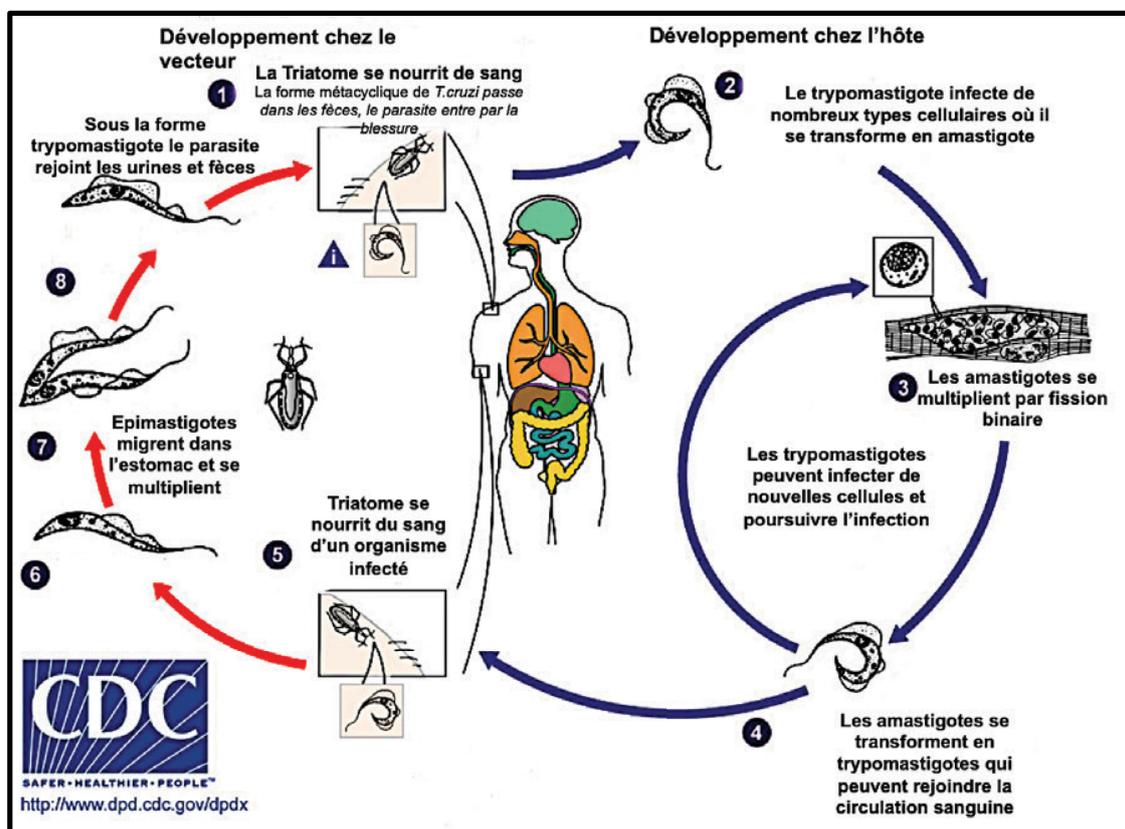


FIGURE 1.IV.1 : CYCLE DE VIE DU PARASITE *TRYPANOSOMA CRUZI* (ISSU DU SITE INTERNET DU CDC)

Le parasite se transmet au triatome lorsque ce dernier se nourrit du sang d'un hôte vertébré infecté (Figure 1.IV.1 (5)). Une fois dans l'estomac, le parasite va effectuer différentes transformations en formes non infectieuses (sphéromastigotes, épimastigotes) qui se multiplient par fission binaire (Figure 1.IV.1 (6) (7)), jusqu'à atteindre le stade infectieux

trypomastigote. A ce stade, le parasite est transféré dans les urines et les excréments du vecteur à travers des mécanismes peu explorés (**Figure 1. IV.1 (8)**).

Lors d'une morsure/piqûre d'un mammifère, comme l'Homme, par le triatome, le parasite est transmis par contact des excréments avec les muqueuses, les blessures cutanées. En effet les fèces de l'insecte hématophage sont déposées à proximité du point de piqûre et l'individu va les mettre au contact de la plaie en se grattant (**Figure 1. IV.1 (1)**). Une fois dans la circulation sanguine, il est capable d'échapper au système immunitaire et de pénétrer les cellules de l'hôte en utilisant des mécanismes que je décrirai à un niveau moléculaire dans la suite de l'introduction. A l'intérieur de la cellule, *T. cruzi* se différencie en forme amastigote non infectieuse pour se multiplier, avant de repasser sous forme trypomastigote capable de sortir de la cellule et de rejoindre la circulation sanguine (**Figure 1. IV.1 (2) (3) et (4)**) (**Coura and Borges-Pereira, 2012**).

IV.2.b. Nouvelles niches écologiques : une menace persistante ?

Même si la présence de *T. cruzi* chez l'Homme remonte à des milliers d'années, leur rencontre est fortuite et relativement rare. L'influence de l'Homme (déforestation, chasse, extension des zones agricoles) a modifié l'environnement en réduisant les sources de nourriture du milieu naturel des triatomes et les a rapprochés des zones domestiques et péri-domestiques. Ceux-ci, attirés par la lumière et les stocks de nourriture, s'introduisent dans les habitations pendant la nuit. Ce changement de niche écologique s'est accompagné par une adaptation des triatomes (réduction de la taille, système sensoriel simplifié, dépendance à l'homme pour se disperser géographiquement par transport passif).

Cette capacité à s'adapter à l'environnement humain pose de grands problèmes aujourd'hui car même si de nombreux moyens ont été trouvés pour limiter la présence des insectes (peinture avec insecticides par exemple), et même si nous sommes capables de contrôler quelques espèces domestiques, rien n'exclut la possibilité d'une nouvelle invasion par la centaine d'espèces sauvages existantes. De même, vu la longévité du parasite chez son hôte, et la centaine d'espèces pouvant être infectées, il existe un réservoir important ne pouvant être éradiqué (**Ayo et al., 2013; Coura and Borges-Pereira, 2012**).

IV.3. Evolution de la réponse immunitaire : de la résistance à la maladie

IV.3.a. Réponse du système immunitaire

La pénétration de *Trypanosoma cruzi* dans la circulation sanguine déclenche rapidement une réponse inflammatoire et la mobilisation des cellules présentatrices d'antigènes (macrophages, cellules dendritiques) dans les premières heures de l'infection à travers la reconnaissance des PAMPs (pathogen associated molecular patterns) à la surface du parasite, notamment par les récepteurs de type Toll à la surface des macrophages et des cellules dendritiques. L'induction du signal s'effectue par l'intermédiaire du domaine Toll/IL-R des récepteurs conduisant à la production d'interleukines et de façon consécutive de cytokines inflammatoires (IFN γ et TNF α sont essentielles au contrôle du parasite) (**Basso, 2013; Machado et al., 2012**). Ces cytokines participent au développement de l'immunité adaptative (prolifération, maturation des lymphocytes B et T, et des cellules NK).

L'effet immunosuppresseur de *T. cruzi* (notamment en induisant la production de cytokines anti-inflammatoires), n'empêche pas une production massive d'anticorps IgG et IgM dans les instants précoces de la phase aigüe. Ces anticorps, dirigés en majorité contre la calréticuline, sont impliqués dans la lyse du parasite en activant les voies du complément. De façon intéressante, le rapport IgG/IgM évolue et détermine si nous sommes dans les instants précoces, intermédiaires ou tardifs de la phase aigüe (**Basso, 2013; Machado et al., 2012**).

IV.3.b. Résistance au système immunitaire

Face à cette réponse immunitaire, *T. cruzi* possède de nombreuses protéines lui permettant de résister aux assauts du système immunitaire et notamment du système du complément.

Les *T. cruzi* Decay accelerating factors (T-DAF), les Complement Regulatory Proteins (CRP), le Complement Receptor 2 Inhibitor (CRIT) inhibent le complément, notamment en interférant avec l'assemblage de la C3 convertase. La cruzipaine, une protéase à la surface du parasite, le protège en dégradant les IgG en fragments Fab' $_2$ et inhibe également la production des cytokines pro-inflammatoires. En parallèle les racémases interconvertissent les énantiomères de L- et D-proline, qui incorporée aux protéines du parasite, lui permet de résister à la protéolyse (**Berasain et al., 2003; Meirelles et al., 1992; Osorio et al., 2012; Scharfstein et al., 2000; Stempin et al., 2002**).

Enfin, la calréticuline de *T. cruzi* est capable d'interagir avec la MBL, la ficoline L et C1q *in vitro* et avec une moindre affinité avec les régions collagènes de C1q seules, et co-localise avec C1q à la surface du parasite *in cellulo* (Ferreira et al., 2004a, 2005; Ramírez et al., 2011a, 2011b; Sosoniuk et al., 2014; Valck et al., 2010). Dans le cas de C1q et de la L-ficoline, cette interaction avec la calréticuline a pour effet d'inhiber les voies classique et lectine du système du complément au niveau de la protéolyse de la molécule C4 en C4b et C4a. Même si le mécanisme d'inhibition n'est pas connu, il se pourrait que la CRT entrave le bon fonctionnement des protéases associées à C1q en créant un encombrement stérique au niveau du site actif de la protéase C1s (Figure 1.IV.2) (Ferreira et al., 2004a, 2005; Ramírez et al., 2011a; Valck et al., 2010). Cette capacité de la calréticuline à inhiber la voie classique du complément pourrait contribuer à l'importante résistance de la forme infectieuse trypanomastigote à la lyse cellulaire, la voie alterne serait donc le seul recours efficace à l'élimination du parasite (si les facteurs B et D sont éliminés, la lyse du trypanomastigote est inhibée).

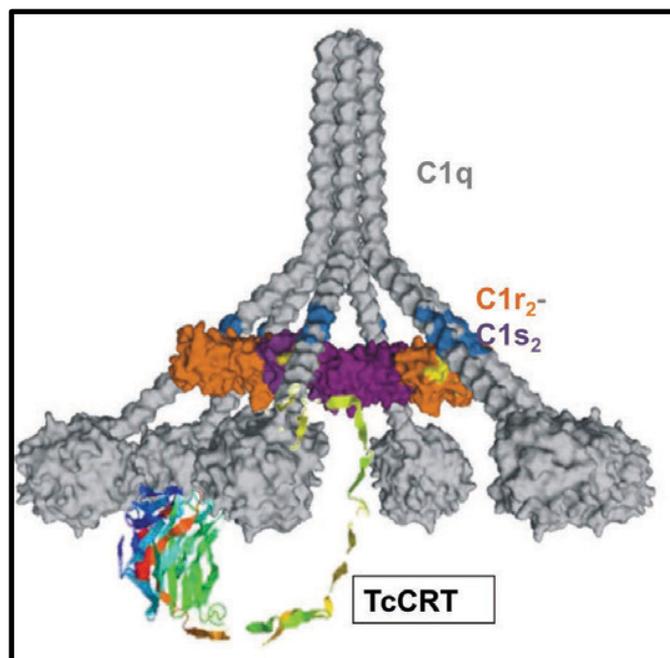


FIGURE 1.IV.2 : PREMIER MODELE D'INHIBITION DE LA VOIE CLASSIQUE DU COMPLEMENT PAR LA TCCRT

La calréticuline interagit avec les domaines globulaires et collagènes de C1q, mais ne peut pas décrocher les protéases déjà fixées. Elle apporterait alors un encombrement stérique au niveau du site actif de la protéase C1s.

Ces observations ont été confirmées *in vivo* en utilisant des souches mutantes. La souche +/- (dans laquelle un des deux allèles du gène codant pour la calréticuline a été supprimé et qui

exprime environ six fois moins la protéine que la souche sauvage) est beaucoup plus sensible à la lyse que la souche sauvage, elle même plus sensible que la souche +/+ (où l'un des allèles a été muté pour surexprimer la calréticuline qui l'exprime deux fois plus). L'effet de la CRT est lié à C1, puisque l'ajout de C1 exogène en excès dans des expériences de lyses *in vitro* rétablit l'élimination du parasite quelle que soit la souche. Cette inhibition peut également être observée en analysant les fèces de *T. infestans* nourris avec du sang contenant les différentes souches : étant donné que l'on retrouve davantage de la souche +/+, on peut supposer qu'elle résiste davantage au complément dans le sang ingéré par le triatome (**Sánchez Valdéz et al., 2013**).

IV.3.c. La réponse immunitaire, acteur majeur de la phase chronique ?

Une forte réponse immunitaire permet l'élimination et le contrôle quasi total du parasite à la fin de la phase aigüe chez la plupart des patients. Cette baisse du nombre de parasites s'accompagne d'une régulation et d'une diminution de la réponse inflammatoire. Les individus infectés entrent alors dans la phase intermédiaire et peuvent y rester jusqu'à la fin de leur vie. Dans certains cas, alors que le niveau du parasite reste très bas, la réponse inflammatoire s'amplifie à nouveau et conduit à la phase chronique.

Les dégâts tissulaires sont principalement dus à l'inflammation (immunologie pathologique), le stress oxydant (déficience des mitochondries, ERO) et la dégradation de la matrice extracellulaire. Les symptômes les plus graves apparaissent notamment au niveau des cellules cardiaques où l'inflammation est accrue en comparaison des autres types tissulaires. Ces observations peuvent s'expliquer par la présence abondante de la calréticuline à la surface des myocytes, où elle intervient dans le développement du muscle cardiaque. De ce fait la calréticuline à la surface de *T. cruzi* entraîne une production massive d'anticorps par le système immunitaire de l'hôte, réagissant avec la calréticuline endogène qui participeront indirectement à la réponse inflammatoire dirigée vers ces tissus (**Basso, 2013; Machado et al., 2012; Michalak et al., 2009**).

IV.4. Invasion par dégradation tissulaire

IV.4.a. Interaction à la surface de la cellule hôte

Trypanosoma cruzi possède à sa surface de nombreuses protéines similaires aux mucines (gp35/50, gp82), des glycoprotéines qui chez l'homme couvrent et protègent les épithéliums intestinaux et respiratoires. Les glycoprotéines comme la gp82 permettent l'adhésion du

parasite à la surface cellulaire en interagissant avec la fibronectine et en activant les voies de mobilisation et de signalisation du calcium. De façon intéressante, *T. cruzi* peut moduler la nature des protéines à sa surface, ce qui lui permet de résister au système immunitaire et d'infecter de nombreux types cellulaires. (Mortara et al., 1992; Osorio et al., 2012; Ouaiissi et al., 1984; Pereira-Chiocola et al., 2000; Ruiz et al., 1998; Schenkman et al., 1993; Turner et al., 2002; Villalta et al., 2008; Yoshida et al., 1989).

Les mucines sont décorées par un acide sialique reconnu par les récepteurs membranaires siglecs (sialic acid-binding immunoglobulin-like lectins) que l'on retrouve à la surface des cellules immunitaires comme les macrophages. Toutefois *T. cruzi* n'est pas capable de synthétiser l'acide sialique, ni son précurseur. Le parasite utilise donc des trans-sialidases, qui catalysent le transfert des acides sialiques, initialement présents à la surface des cellules hôtes, sur les glycoprotéines à la surface du parasite. Ce transfert est crucial à la viabilité et l'infection de *T. cruzi* (Figure 1.IV.3 (2) et (3)) (Osorio et al., 2012; Villalta et al., 2008).

IV.4.b. Remodelage et destruction de la matrice extracellulaire

Après son adhésion aux cellules, la forme infectieuse trypanostigote de *T. cruzi* module et détruit la matrice extracellulaire. Trois protéines de l'hôte participent à la fixation de *T. cruzi* à la matrice extracellulaire: la thrombospondine, la laminine γ et la galectine 3, dont la surexpression des gènes est déclenchée par la protéine gp83 du parasite.

La laminine γ (codée par *LAMC1*), interagit avec la laminine des cellules musculaires et la galectine 3, une lectine humaine formant un pont entre les mucines exprimées à la surface de *T. cruzi* et le domaine carbohydrate de la laminine. Elle peut également fixer de nombreuses protéines de la matrice extracellulaire (collagène IV, fibronectine, elastine, tenascine...) (Cardenas et al., 2010; Kleshchenko et al., 2004; Nde et al., 2006, 2010, 2012).

La calréticuline de *T. cruzi* interagit avec la thrombospondine 1 (codée par le gène *THSBI*). Cette dernière se fixe à l'héparine, aux collagènes de la matrice extracellulaire et également à de nombreuses protéines de la famille de la laminine (Cardenas et al., 2010; Nde et al., 2010, 2012; Simmons et al., 2006).

D'autres protéines interagissent avec le cytosquelette : la pénétrine, DAG1 (crée des liens transmembranaires entre le cytosquelette et la matrice extracellulaire), MYOC (rôle dans la fonction du cytosquelette, transport des vésicules et conformation de la matrice) (Osorio et al., 2012; Santana et al., 1997; Villalta et al., 2008).

Une fois ancré à la matrice, le parasite la dégrade et progresse en force entre les cellules grâce à de nombreuses protéases (cruzipaine, MMP2 (Matrix Metallo Protéinase 2 et la prolyl oligopeptidase (Tc80)), toutes spécifiques des collagènes de type I et IV.

IV.5. Mécanismes d'invasion non destructifs

Au cours de l'évolution, *T. cruzi* a également développé des mécanismes de pénétration cellulaire plus subtils, en détournant de nombreux mécanismes de l'hôte, favorisant son entrée dans les lysosomes et macrophages où il est à l'abri du système immunitaire à travers une stratégie de type « Cheval de Troie ».

IV.5.a. Le modèle dépendant des lysosomes

Ce premier modèle de pénétration fait intervenir un phénomène de réparation naturel des membranes détourné par le parasite à son avantage. *T. cruzi*, une fois attaché à la cellule, utilise des seconds messagers comme le calcium, pour recruter des vésicules et se différencier comme décrit sur la **Figure 1.IV.3**.

La cruzipaine (CZP) et l'oligopeptidase B (OPB) déclenchent la mobilisation du calcium en générant la bradykinine et le 1,4,5-inositol triphosphate (**Figure 1.IV.3 (1)**). Le calcium module l'activité motrice de la kinésine pour former la vacuole lysosomale dont le parasite devra s'échapper grâce à l'activité des trans-sialidases, et d'enzymes le protégeant des ERO (hémolysine TcTox, tryparedoxine peroxydase TcMPX/TcCPX/TcAPX) (**Figure 1.IV.3 (4) et (5)**).

Une fois entré dans la cellule de manière passive, la protéine phosphatase 2A (PP2A), la PPLC, et le protéasome interviennent dans la différenciation de la forme trypomastigote en amastigote (**Figure 1.IV.3 (6) et (7)**), tandis que la calcineurine assure la thermorésistance membranaire du parasite (**Araya et al., 2008; Kaplan et al., 1998; Osorio et al., 2012; Villalta et al., 2008**).

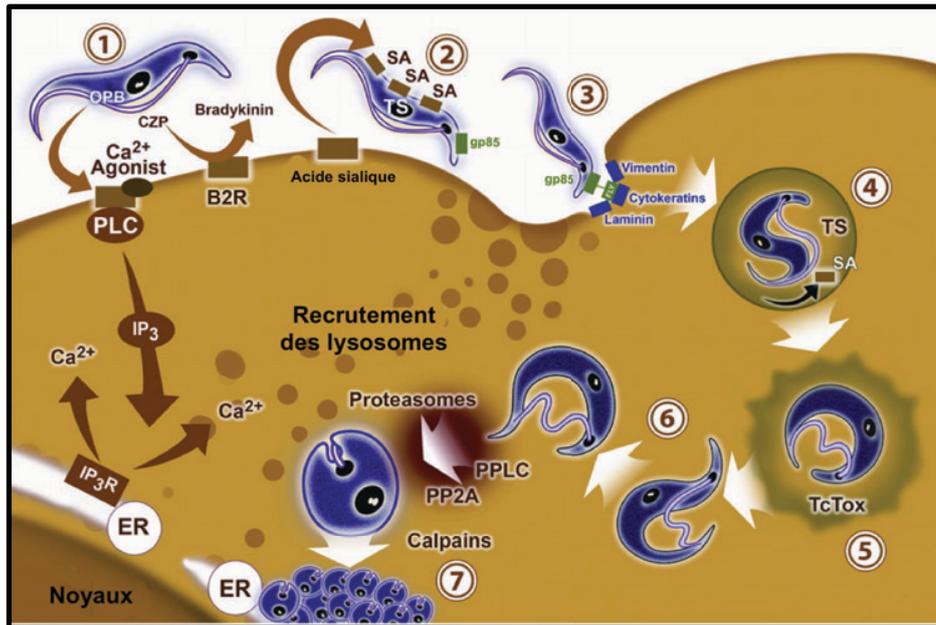


FIGURE 1.IV.3 : MODELE DE PENETRATION DE *T. CRUZI* DEPENDANT DES LYSOSOMES (OSORIO ET AL., 2012)

IV.5.b Le modèle phagocytose : implication de la calréticuline

Ce second mécanisme captivant fait intervenir la protéine C1q. En effet, C1q purifié ou du sérum mis au contact du parasite favorise la pénétration de la forme trypomastigote, spécifiquement, dans les macrophages, monocytes et fibroblastes. Le trypomastigote est également capable de fixer C1q avec un rendement mille fois supérieur à la fixation de C1q sur les épimastigotes (Rimoldi et al., 1988, 1989).

De nombreux antigènes, ainsi que la composition de la membrane en protéines et lipides de *T. cruzi*, ont été caractérisés, permettant d'établir une liste potentielle des molécules potentiellement reconnues par C1q à la surface du parasite (Aslund et al., 1994; Bontempi et al., 1993; Campetella et al., 1992; Kaplan et al., 1998; Krautz et al., 1998; Quiñones et al., 2004; Swindle et al., 1988).

Une de ces protéines, déjà décrite comme intervenant dans la résistance du parasite au système immunitaire et à son adhésion à la matrice extracellulaire, a très vite été proposée comme un acteur important de la stratégie mise en place par le parasite lors de l'infection, puisque déjà identifiée comme un récepteur de C1q à la surface des cellules de mammifères (Sim et al., 1998): la calréticuline.

IV.6. Rôle de l'interaction C1q/calréticuline dans le modèle de phagocytose

L'importance de la calréticuline dans l'infectivité du parasite a été confirmée *in vivo* à travers l'utilisation des souches de *T. cruzi* -/+ et *T. cruzi* +/+. En effet, aucun parasite n'est détectable après une inoculation de la souche +/- dans des souris immunocompétentes alors que c'est le cas pour les autres (WT et +/+ ont le même comportement) (**Sánchez-Valdéz et al., 2014**).

De plus, le taux d'ARN messagers codant pour la calréticuline augmente drastiquement lors des premières étapes de l'infection (**Ramírez et al., 2011b, 2012**).

IV.6.a. Organisation modulaire et localisation de la calréticuline

La calréticuline de *T. cruzi*, identifiée pour la première fois en 1991 (**Ramos et al., 1991**), possède 41% d'identité de séquence avec les calréticulines de mammifère (**Aguillón et al., 2000**). Elle conserve les motifs consensus et une organisation modulaire comme décrit dans les différentes calréticulines: un domaine globulaire N-terminal, un domaine P riche en prolines et flexible, et une hélice C-terminale riche en résidus acides. Elle ne possède que 2 des 3 cystéines aux positions consensus (**Ferreira et al., 2004b, 2005**).

Initialement, la CRT est présente de manière abondante dans le réticulum endoplasmique des formes épimastigote et trypomastigote, où elle joue un rôle de protéine chaperonne (**Souto-Padrón et al., 2004**). *T. cruzi* possède toutes les protéines nécessaires à la fonction chaperonne dans le RE, à l'exception de la calnexine (**Caramelo et al., 2004; Conte et al., 2003; Labriola et al., 1999, 2011; Sánchez Valdéz et al., 2013; Trombetta et al., 2001**).

Il existe néanmoins une grande différence dans le cycle de la fonction chaperonne de *T. cruzi* : *T. cruzi* est incapable de synthétiser l'intermédiaire Glc₃Man₉GlcNAc₂. De ce fait, le substrat Glc₁Man_{7,8,9}GlcNAc₂ reconnu par la CRT ne résulte pas d'une déglycosylation partielle comme chez la plupart des eucaryotes (**Conte et al., 2003; Labriola et al., 2011**).

Toutefois, alors que la calréticuline n'est observée dans la forme épimastigote qu'à l'intérieur du cytoplasme, du réservoir, des kinétoplastes, elle est également exportée à la surface de la forme infectieuse trypomastigote, confirmant son rôle majeur dans l'infectivité déjà évoqué (adhésion à la matrice, pénétration dans les cellules, résistance au complément) (**Souto-Padrón et al., 2004**).

IV.6.b. Mécanisme d'infectivité : rôle des anticorps

Paradoxalement, les tentatives de traitement telles que des vaccins, en immunisant les sujets avec de la calréticuline recombinante, stimulent fortement la pénétration du parasite dans les cellules hôtes. Vraisemblablement C1q reconnaît le domaine CH2 de la partie Fc des anticorps (notamment des IgG1) qui constituent de nouveaux récepteurs à la surface du parasite (Aguilar et al., 2005; Ferreira et al., 2005; Ramírez et al., 2011a, 2011b; Xu et al., 1994).

Ces résultats ont ouvert de nouvelles perspectives thérapeutiques, puisque la génération d'anticorps Fab'₂ (dont le domaine Fc a été digéré) permet de masquer la calréticuline sans créer de nouveau site de fixation pour C1q. Il a d'ailleurs été constaté qu'une pré-incubation de la calréticuline ou du parasite avec ces fragments d'anticorps diminue de 70 à 80% le recrutement de C1q à la surface du parasite et l'interaction entre les protéines. L'immunisation de souris avec ces fragments préserve également leur survie face à l'infection (Aguilar et al., 2005; Ramírez et al., 2011a, 2011b, 2012)

Ceci est d'autant plus important que la génération d'IgG anti-calréticuline, qu'ils soient dus à l'infection ou à une immunisation, engendre à terme la synthèse d'IgG dirigés contre la calréticuline humaine et induit les lésions à l'origine de la phase chronique de la maladie de Chagas (Ramírez et al., 2011b; Ribeiro et al., 2009).

Dans ce contexte, l'utilisation pour l'immunisation d'un patient d'une souche de *T. cruzi* -/+ deviendrait très intéressante pour générer une immunoprotection sans déclencher la production d'anticorps anti-CRT pouvant réagir avec la calréticuline endogène et entraîner des dégâts (Sánchez-Valdéz et al., 2014).

IV.6.c. Mécanisme d'infection : un mimétisme avec les cellules apoptotiques ?

Malgré la description de l'importance de l'interaction entre la protéique C1q et la calréticuline de *T. cruzi* ou les anticorps anti-calréticuline générés pendant l'infection, pour pénétrer dans les macrophages et les cellules de l'hôte (Figure 1.IV.4 (a) et (d)), le mécanisme par lequel C1q favorise l'entrée dans la cellule, et notamment la présence d'un récepteur, est encore mal compris. Nos collaborateurs chiliens se sont également intéressés à un problème de santé publique, celui de la transmission du parasite *T. cruzi* de la mère à l'enfant durant la grossesse à travers le placenta. Ils ont en particulier analysé le rôle de C1q et des calréticulines du

parasite et de l'hôte dans ce contexte. Comme illustré sur la **Figure 1.IV.4**, ces trois protéines semblent jouer un rôle important. C1q permet de mettre en contact le parasite et la cellule à infecter *via* l'interaction avec les calréticulines. Une fois à son contact, le parasite est engouffré de manière passive dans la cellule, comme par exemple dans le cas du macrophage, décrit **figure 1.IV.5**, de manière similaire aux cellules apoptotiques.

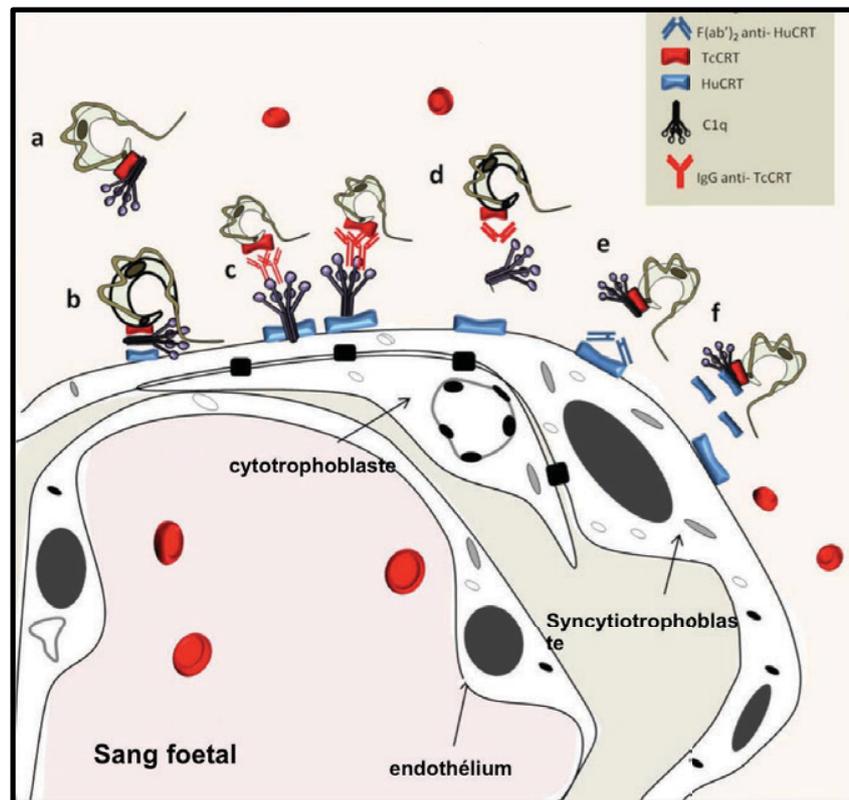


FIGURE 1.IV.4 : MECANISME D'INFECTION DE *T. CRUZI* : EXEMPLE DES CELLULES DU PLACENTA (CASTILLO ET AL., 2013)

Lors de la transmission du parasite de la mère à l'enfant, la protéine C1q d'origine maternelle favorise la pénétration du parasite dans les cellules fœtales, de même que les anticorps dirigés contre la calréticuline de *T. cruzi* comme j'ai déjà pu le présenter ((b) et (c)). Des fragments Fab'₂ d'immunoglobulines dirigés contre la calréticuline de *T. cruzi* bloquent la pénétration du parasite. Ici la nouveauté est la participation de la calréticuline de la cellule infectée. En effet l'immunisation par des fragments Fab'₂ d'immunoglobulines dirigés contre la calréticuline humaine, présente à la surface des cellules fœtales, inhibe la pénétration du parasite, de la même manière que la compétition par ajout de calréticuline humaine recombinante à C1q en phase liquide ((e) et (f)). Il est donc proposé que C1q puisse reconnaître la calréticuline du parasite ou les anticorps à sa surface, et alors jouer le rôle de molécule pontante en se fixant à la calréticuline de l'hôte pour mettre au contact le parasite et la cellule à infecter.

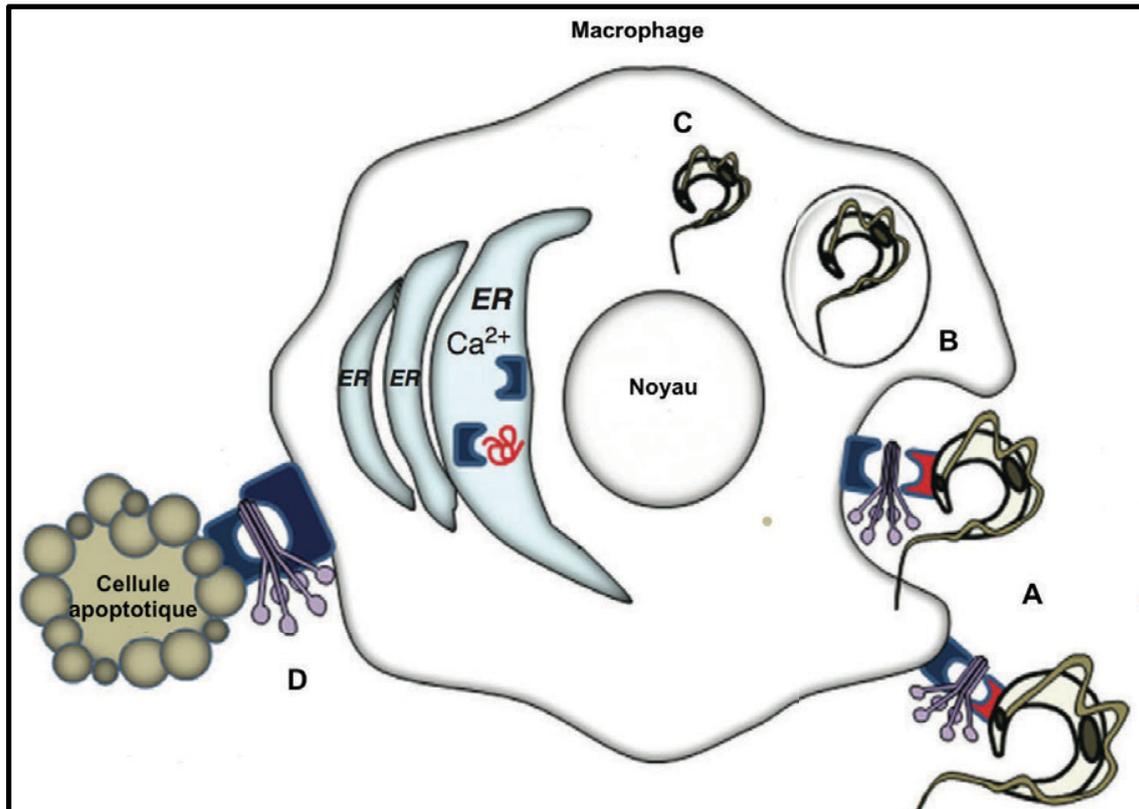


FIGURE 1.IV.5 : MECANISME D'INVASION DE TYPE « CHEVAL DE TROIE » DE *T. CRUZI* DANS LES MACROPHAGES (RAMIREZ ET AL., 2011B)

Outre l'importance de C1q et de la calréticuline de *T. cruzi* pour favoriser l'entrée du parasite dans les macrophages, les études sur le placenta ont montré la participation de la calréticuline de l'hôte. C1q favoriserait alors la pénétration du parasite dans les macrophages, en jouant le rôle de molécule pontante avec la calréticuline de l'hôte à leur surface, comme décrit pour les cellules apoptotiques (**D**). Contrairement aux cellules du placenta, il ne s'agit plus ici d'une simple mise en contact à la surface de la cellule, mais l'interaction déclenche le mécanisme de phagocytose permettant au parasite d'être internalisé dans une vacuole phagolysosomale (**B**), plutôt que de détruire les tissus. Enfin, *T. cruzi* peut s'évader de cette vacuole pour se multiplier et infecter d'autres types cellulaires (**C**).

V. Rôle de la calréticuline dans la phagocytose: l'exemple de la calréticuline d'*Entamoeba histolytica*

V.1. *Entamoeba histolytica*, parasite responsable de l'amibiase

Entamoeba histolytica est un des agents pathogènes causant l'amibiase. Même si l'apparition de diarrhées est notée chez des individus infectés par la forme non invasive du parasite, cette infection est le plus souvent asymptomatique. Néanmoins dans certains cas, les symptômes provoqués par la forme invasive de l'amibe sont caractérisés par des colites et des déséquilibres gastro-intestinaux provoqués principalement par la dégradation de la muqueuse épithéliale de l'intestin. Dans des situations extrêmes, la dégradation tissulaire provoquée par le parasite peut s'étendre au foie, aux poumons ou au cerveau et provoquer des abcès. Il est aujourd'hui estimé que la maladie touche 50 millions de personnes, et est responsable de 40 000 à 100 000 morts par an, notamment au Mexique, Amérique du Sud, Asie, Afrique et dans certaines îles du Pacifique.

L'infection se produit lors de l'ingestion d'eau et de nourriture contaminées par des kystes de l'amibe (et dans des cas très rares par transmission sanguine lors de rapport sexuels) (**Figure 1.V.1 (2)**). Cette forme résistante du parasite est protégée de l'acidité de l'estomac grâce à une paroi riche en chitine. Contrairement à *T. cruzi*, le cycle de vie de *E. histolytica* ne nécessite la présence que d'un seul hôte. Une fois ingérés, les kystes transitent dans l'intestin grêle où ils libèrent leurs 4 noyaux à travers des mécanismes inconnus. Ces noyaux se divisent par fission binaire et s'entourent d'un cytoplasme pour donner naissance à huit trophozoïtes, dotés d'un flagelle et capables de se déplacer jusqu'au lumen du côlon. Au stade trophozoïte, l'amibe peut donner de nouveaux kystes ou se diviser. Elle se présente sous deux formes (**Figure 1.V.1 (3), (4) et (5)**). :

- la forme *histolytica minuta* vit dans le lumen ou parfois à la surface de la muqueuse du côlon. Sous cette forme, l'amibe se nourrit essentiellement de bactéries présentes dans l'intestin et ne provoque aucun symptôme. Elle peut également reformer des kystes qui seront éliminés dans les selles (**Figure 1.V.1 (A), (1)**)
- la forme *histolytica histolytica* est la forme à l'origine de la maladie. Elle pénètre la muqueuse du côlon par effraction et dégrade les tissus de l'intestin par des mécanismes que j'aborderai dans les prochaines sections. La possibilité d'évoluer sous cette forme dépend de l'influence des génomes de l'hôte et du parasite, ainsi que des

modifications de l'environnement comme la réponse du système immunitaire ou l'évolution de la flore bactérienne (**Figure 1.V.1 (B), (C)**).

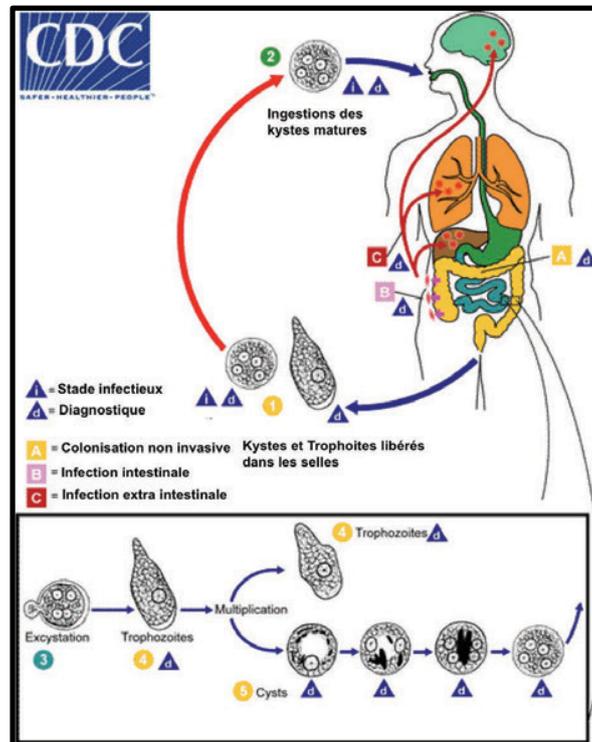


FIGURE 1.V.1 : CYCLE DE VIE DE *ENTAMOEBIA HISTOLYTICA* (ISSU DU SITE DU CDC)
Description du cycle de vie des formes kystes et trophozoïtes de l'amibe pathogène *Entamoeba histolytica*. Le code couleur correspond aux couleurs utilisées dans le texte.

La fréquence importante de la forme asymptomatique de l'infection montre que la virulence est accessoire à la réplication, la transmission et la survie du parasite et ne lui confère aucun avantage de sélection. Il a d'abord été proposé que la capacité de *Entamoeba histolytica* à phagocyter des bactéries pour se nourrir était liée à la virulence, mais il a été récemment montré que la souche non pathogène *Entamoeba dispar*, longtemps considérée comme la forme non invasive de *histolytica*, était également capable de se nourrir de bactéries. Il semblerait que l'évolution vers la forme invasive pourrait être provoquée par une privation de nourriture en influençant directement le génome du parasite (recombinaison homologue, repression/activation de gènes, transfert latéral d'ADN bactérien). *E. histolytica* a en effet un métabolisme très réduit et doit trouver des nutriments dans des sources exogènes, notamment des acides aminés, acides nucléiques et lipides. Cette amibe ne possède pas les voies de synthèse des acides aminés (exceptées la cystéine et la sérine), des bases (purine et pyrimidine) ainsi que des acides gras, et il a été montré que les cultures en milieu anoxique nécessitaient l'évolution du parasite vers une forme trophozoïte capable de phagocyter pour

obtenir les nutriments manquants par pinocytose. Au contraire, *E. dispar* possède un métabolisme associé au commensalisme et est très difficilement cultivable en milieu anoxique. De ce fait l'étude des différences entre les génomes très similaires de ces deux souches (l'identité des ARN est de 98%) peut apporter des réponses importantes sur les gènes codant pour des effecteurs essentiels à la virulence (**Marie and Petri, 2014; Ralston and Petri, 2011**).

V.2. Adhésion à la surface des cellules

Une lectine spécifique du galactose (Gal) et des N-acetyl-D-galactosamine (GalNac), présente uniquement à la surface de la forme trophozoïte de *E. histolytica*, peut interagir avec les mucines et les cellules épithéliales intestinales de l'hôte. Même si d'autres protéines participent à l'adhésion, cette lectine semble cruciale et, lors d'une infection, 80% des anticorps générés sont dirigés contre cette dernière. Un autre domaine de la protéine contient un site de fixation à la fibronectine, participant également à l'adhésion (**Marie and Petri, 2014; Petri et al., 2002; Ravdin and Guerrant, 1981**).

Cette étape d'adhésion aux cellules est le premier niveau auquel l'amibe peut être maîtrisée. Initialement, l'entrée du parasite provoque un épaississement de la couche de mucines pour éviter qu'il n'entre en contact avec les cellules épithéliales du côlon. En effet, des modèles d'étude dépourvus de mucines sont drastiquement plus sensibles à l'infection que les modèles sauvages. De même, la production de nombreux anticorps (IgA) contre la lectine de surface permet de limiter la fixation du parasite. Ces régulations permettent de maintenir l'infection sous sa forme non invasive dans 15% des cas (**Abd-Alla et al., 2006; Haque et al., 2001, 2006; Prathap and Gilman, 1970**) (**Figure 1.V.2 (A)**).

V.3. Survie à l'interface hôte-pathogène et évolution vers la maladie

L'adhésion de la forme invasive aux cellules épithéliales du côlon ne suffit pas à déclencher la maladie. Cette dernière est initiée par la destruction des tissus et de nombreux facteurs et mécanismes, décrits sur la **figure 1.V.2**, influent sur la balance permettant de contrôler la virulence et de faire évoluer l'infection.

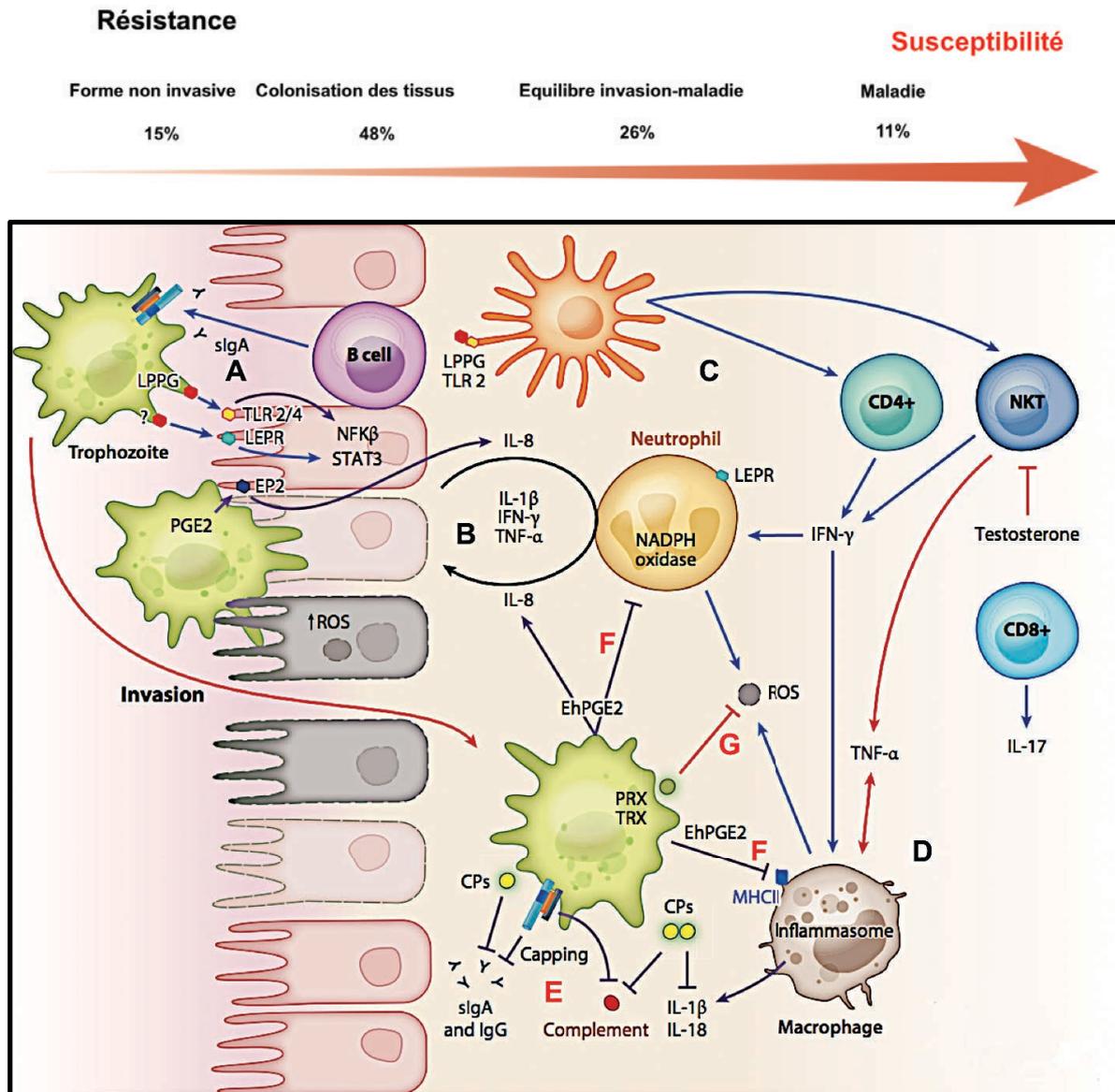


FIGURE 1.V.2 : REPOSE IMMUNITAIRE LORS DE LA COLONISATION DES TISSUS (MARIE AND PETRI, 2014)

Si les anticorps empêchant l'amibe d'adhérer aux cellules sont dépassés (A), la réponse inflammatoire induite par les neutrophiles et les macrophages maintiennent l'amibe au niveau de l'épithélium cellulaire (B). Si elle pénètre plus profondément dans les tissus, le système du complément, les macrophages et les cellules de l'immunité adaptatives prennent le relais (C et D). En réponse au système immunitaire, *E. histolytica* possède les outils pour interférer avec la réponse des neutrophiles et la maturation des macrophages (F), la génération des espèces réactives de l'oxygène (G), et dégrader les composants du complément (E). Le système immunitaire stoppe la progression de l'amibe dans le lumen dans 60% des cas, alors que l'amibe prend le dessus dans 11% des infections conduisant à la maladie.

V.3.a. Contact avec les tissus et mise en place de la réponse immunitaire

Entamoeba histolytica sécrète de nombreuses glycosidases et protéases lui permettant de dégrader les mucines et atteindre les cellules épithéliales et de coloniser les tissus (dans 48% des cas). Ce contact avec les tissus génère une réponse immunitaire importante, maintenant un équilibre entre colonisation et maladie (26% des cas) (**Kissoon-Singh et al., 2013**). La reconnaissance de la lectine de surface par les récepteurs TLR2 et 4 à la surface des cellules épithéliales active par exemple la voie de NF- κ B et la production concomitante de nombreuses cytokines inflammatoires ainsi que le recrutement des neutrophiles. La réponse inflammatoire rompt les jonctions intracellulaires facilitant l'intervention des neutrophiles, dont le rôle en plus de maintenir la réponse inflammatoire est de dégrader le parasite en générant des ERO (**Figure 1.V.2 (B)**) (**Bansal et al., 2009; Chadee et al., 1987; Espinosa-Cantellano and Martínez-Palomo, 2000; Galván-Moroyoqui et al., 2011; Prathap and Gilman, 1970**).

Si la barrière de cellules épithéliales est franchie et que l'amibe atteint la lamina, elle est reconnue par les récepteurs TLR-2 des cellules dendritiques, qui vont activer les cellules NK et les lymphocytes T CD4⁺ et moins fréquemment les macrophages. Il a été montré que la maîtrise de l'infection était principalement associée à la production de IFN- γ et d'interleukine 17, alors que la production de TNF α augmente l'inflammation des tissus (**Guo et al., 2009; Haque et al., 2007; Helk et al., 2013; Lotter et al., 2009; Peterson et al., 2010**) (**Figure 1.V.2 (C) et (D)**). Les antigènes à la surface de l'amibe sont également responsables de la production d'IgG nécessaire à l'activation de la voie classique du complément, également actrice de la protection immunitaire.

La réponse immunitaire semble également affectée par l'environnement. Un lien a en effet été établi entre le développement de la forme virulente de la maladie et les régions souffrant de malnutrition. Les enfants en carence ont en effet un faible niveau de leptine, la protéine de la satiété, qui serait directement reliée à la susceptibilité de développer la maladie (**Duggal et al., 2011; Faggioni et al., 2001; Mondal et al., 2009**) en agissant directement sur le système immunitaire. Les leptines réguleraient le système immunitaire, en augmentant la synthèse de peptide antimicrobien (**Rajala et al., 2014**), prévenant l'apoptose, augmentant la sécrétion des mucines de protection et améliorant la réparation et la prolifération des cellules épithéliales (**El Homsy et al., 2007; Rouet-Benzineb et al., 2004; Sukhotnik et al., 2006**).

V.3.b. Evasion du système immunitaire

V.3.b.1 Inhibition du système du complément

Afin de résister au système immunitaire et pour faire basculer l'équilibre à l'interface de l'épithélium vers l'invasion et la maladie comme cela se produit dans 11% des infections, *Entamoeba histolytica* possède de nombreux mécanismes lui permettant de s'évader.

E. histolytica possède un nombre important de protéases à cystéines (environ une cinquantaine) qui, en plus d'intervenir dans la dégradation des mucines et des tissus pour faciliter la pénétration du parasite, sont également dirigées contre des acteurs de la réponse immunitaire. Par exemple la protéase à cystéine 1 (CP1) cible la protéine du complément C3, et également l'interleukine (IL-8), alors que l'IL-18, les IgA, C3 et C5a sont dégradés par CP4. Enfin une dernière protéase, CP5, est responsable de l'élimination des IgG. Toutes ces protéases sont associées à des CP-binding protéines (CPBF) pour assurer leur activité. *E. histolytica* produit également un inhibiteur du chimiotactisme : le MLIF (monocyte locomotion inhibitor factor) (Kelsall and Ravdin, 1993; Que and Reed, 2000; Que et al., 2003; Reed et al., 1995; Rico et al., 2003; Tran et al., 1998).

Hormis la dégradation protéolytique, l'amibe se protège du système du complément en inhibant la fixation des anticorps grâce à la mise en place d'une enveloppe de surface (capping). La lectine Gal/NAc inhibe l'association du complexe d'attaque membranaire (MAC) de même que l'EhGL, adhésine possédant une région similaire à CD59 (Figure 1.V.2 (E)) (Braga et al., 1992; Chávez-Munguía et al., 2012; Ximénez et al., 2014).

Comme de nombreux parasites, il a été montré que la calréticuline des deux souches de *E. histolytica*, virulente et avirulente, était capable d'interagir avec la protéine C1q et d'inhiber la voie classique du système du complément. Pareillement, la présence de cette protéine dans un sérum protège significativement l'amibe de la lyse dépendante du système du complément. Cependant, le contact avec C1q ne se produit qu'en cas d'invasion des tissus par les trophozoïtes, qui présentent à leur surface la CRT (Ximénez et al., 2014).

V.3.b.2 Régulation de l'inflammation et destruction des cellules

E. histolytica peut également s'attaquer directement aux cellules du système immunitaire. La sécrétion d'une prostaglandine (PGE2) inhibe la maturation du complexe majeur d'histocompatibilité II (MHCII) à la surface des macrophages et diminue la production

d'ERO en agissant directement sur la NADPH oxydase à la surface des neutrophiles (**Wang and Chadee, 1995**). Plus radicalement, le parasite peut tuer et internaliser les cellules du système immunitaire, de la même manière que les bactéries, ou les cellules épithéliales (**Ralston and Petri, 2011; Ravdin et al., 1980**) (**Figure 1.V.2 (F)**).

Que ce soit pour faire face aux ERO sécrétés par les macrophages et neutrophiles, ou lors de l'ingestion de ces mêmes cellules, l'amibe est constamment menacée par l'oxydation et doit y faire face. Pour cela *E. histolytica* a la capacité de réduire la quantité d'oxygène, première source de ERO, en transformant la molécule O₂ en eau *via* différentes enzymes. Cette diminution de O₂, couplée à la réduction des stocks d'arginine (intervention d'une arginase) diminue l'activité des NADPH oxydase pour générer le NO (**Band and Cirrito, 1979; Bruchhaus and Tannich, 1994; Choi et al., 2005; Elnekave et al., 2003; Loftus et al., 2005; Ramos-Martínez et al., 2009**). L'amibe possède également à sa surface une thioredoxine et une peroxyrédoxine (Eh29) associée à la lectine Gal/NAc et intervenant dans la réduction des ERO en molécules inertes. En se nourrissant de certaines bactéries, elle s'approprie également des chélateurs de ERO (**Figure 1.V.2 (G)**).

V.4. Elimination des cellules

E. histolytica possède la surprenante capacité de pouvoir tuer différents types de cellules, en induisant leur apoptose par contact physique, puis en les phagocytant à la manière d'un macrophage. Comme je l'ai déjà mentionné, cette compétence est nécessaire pour la survie du parasite (en phagocytant des bactéries dans le lumen du côlon, ou des cellules épithéliales pour obtenir des nutriments, des cellules du système immunitaire pour se protéger, et les cellules de la paroi intestinale pour progresser dans les tissus). Trois grands mécanismes ont été décrits pour assurer cette fonction :

- Le premier mécanisme, le plus couramment décrit et étudié, implique **une cyto-toxicité induite par contact suivie d'une phagocytose**. Les mécanismes par lesquels le contact avec l'amibe induit l'apoptose de la cellule de l'hôte ne sont pas encore totalement décryptés. Le contact avec l'amibe provoque une augmentation importante de la concentration de Ca²⁺, cependant cette variation ne provoque pas à elle seule la mort cellulaire. Le contact induit également la production massive d'ERO dans la cellule et la fragilisation de la membrane plasmique (en faisant intervenir une amoebase, protéine capable de former des pores dans la membrane lipidique et des

protéines de la famille des saposin-like protéines SAPLIPs associées aux lipides). Finalement le contact induit l'apoptose à travers un nouveau mécanisme dépendant de la caspase 3, mais indépendant des caspases 8, 9 et de Bcl2 (**Figure 1.V.3 (A)**) (**Huston et al., 2000; Ragland et al., 1994; Ralston and Petri, 2011; Ravdin et al., 1982, 1985, 1988; Seydel and Stanley, 1998; Yan and Stanley, 2001**).

La mort cellulaire se traduit ici comme attendu par des marqueurs de l'apoptose (fragmentation de l'ADN, exposition de la PS) et est suivie par l'élimination de la cellule apoptotique à travers un mécanisme de phagocytose. L'interaction amibe/cellules induit une signalisation conduisant au remodelage du cytosquelette d'actine de l'amibe pour engouffrer la cellule à éliminer. Ce réarrangement est contrôlé par les protéines G et la myosine de l'amibe (**McCoy and Mann, 2005; Nayal et al., 2004; Somlata et al., 2011; Vines et al., 1998**).

De manière intéressante de nombreuses protéines du réticulum endoplasmique se retrouvent à la surface cellulaire au moment de la phagocytose (**Basu et al., 2001**) comme la calréticuline associée à la lectine Gal/NAc. Ici la CRT pourrait encore une fois, à l'instar de *T. cruzi*, permettre à l'amibe de mimer nos macrophages et utiliser notre propre système immunitaire (C1q) pour favoriser la phagocytose de nos propres cellules (**Basu et al., 2001; Desjardins, 2003; Gagnon et al., 2002; Marion et al., 2005; McCoy and Mann, 2005; Vandivier et al., 2002**).

- Un second mécanisme, très récemment introduit, est la **Trogocytose**. Ce processus, proche du précédent, nécessite également un contact cellulaire et la présence de C2PK et provoque une augmentation de la concentration de Ca^{+} . Toutefois l'amibe phagocyte de très petits fragments de la cellule toujours vivante, provoquant une rupture de la membrane plasmique et une mort très rapide de la cellule. Contrairement au cas précédemment décrit, la cellule une fois morte n'est pas ingérée (**Prathap and Gilman, 1970; Ralston et al., 2014**) (**Figure 1.V.3 (B)**).
- Le dernier mécanisme n'implique aucun contact entre l'amibe et la cellule. Ici le parasite sécrète différents effecteurs pour atteindre la cellule cible à distance (prostaglandine 2 (PGE2) ou la claudine-4). Ces molécules déstabilisent les jonctions intercellulaires en modifiant la concentration intracellulaire des ions sodium et chlorure (**Lejeune et al., 2011; McGowan et al., 1983, 1990**) (**Figure 1.V.3 (C)**).

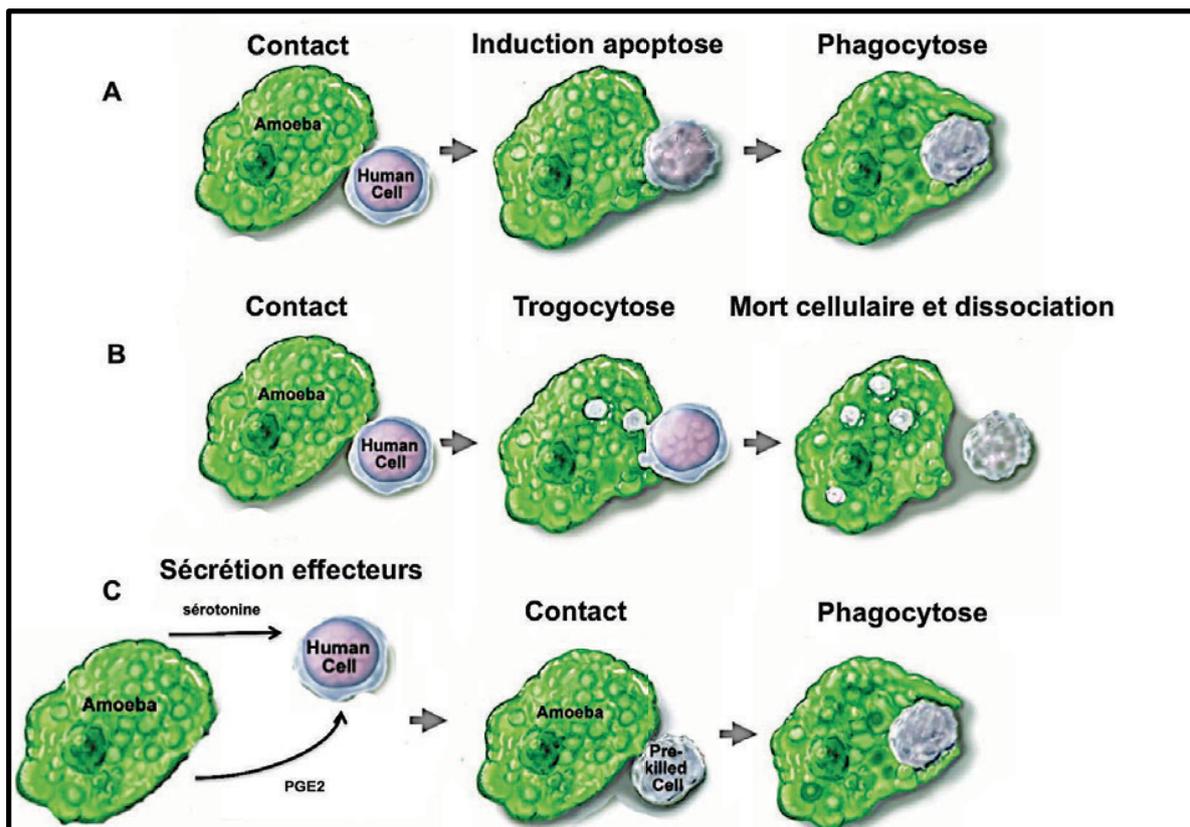


FIGURE 1.V.3 : ILLUSTRATION DES MECANISMES UTILISES PAR *E. HISTOLYTICA* POUR DETRUIRE LES CELLULES. ADAPTE DE (MARIE AND PETRI, 2014; RALSTON ET AL., 2014)

V.5. Rôle de l'interaction C1q/calréticuline dans la phagocytose par l'amibe

V.5.a. Identification, séquençage et localisation de la calréticuline EhCRT

La calréticuline de *Entamoeba histolytica* EhCRT a été isolée et séquencée pour la première fois en 2002 à partir de la souche non virulente HM1:IMSS. Elle a été identifiée en utilisant des anticorps IgA et IgG purifiés à partir de la salive de patients atteints d'amibiase (González et al., 2002). Cette dernière présente 52% d'identité avec la calréticuline humaine, 98% avec la calréticuline de l'autre souche non invasive *Entamoeba dispar*. Nous retrouvons la même organisation modulaire que celle que j'ai décrite jusqu'à présent, le domaine globulaire N-terminal, le domaine P riche en proline et flexible, et l'hélice C-terminale riche en résidus acides. Contrairement à *T. cruzi*, elle possède, comme les calréticulines de mammifères, les 3 cystéines aux positions consensus.

Entamoeba histolytica ne possède pas un réticulum endoplasmique structuré, celui-ci est composé de nombreuses vésicules réparties à proximité de la membrane plasmique interne de l'amibe. Comme attendu, la calréticuline est localisée dans ces vésicules en compagnie d'autres protéines spécifiques du RE comme la protéine homologue de BiP. Elle y joue le rôle de protéine chaperonne à travers sa fonction lectine. Comme *T. cruzi* et à la différence des mammifères, *Entamoeba histolytica* ne possède pas d'homologue de la calnexine et le glycane précurseur Man₅GlcNAc₂ nécessite également l'ajout d'un résidu glucose. **(Banerjee et al., 2007; González et al., 2002).**

V.5.b. Rôle de la EhCRT à la cupule phagocytaire

La forme infectieuse et invasive d'*Entamoeba histolytica* n'exprime pas en permanence la CRT à sa surface. Cette dernière sera exportée à la surface de l'amibe au niveau de la cupule phagocytaire au moment de l'initiation de la phagocytose, où elle interagit avec la protéine C1q (**Figure 1.V.4 (B) et (C)**) **(Ximénez et al., 2014)**. Il avait déjà été préalablement montré que, comme chez les mammifères, C1q ou la MBL étaient des chémoattractants de l'amibe et favorisaient la phagocytose des cellules de l'hôte. Comme l'on sait que le contact entre l'amibe et la cellule hôte induit l'apoptose de cette dernière entraînant l'exposition de signaux eat-me, il se pourrait que la CRT de l'hôte soit également exposée à la surface **(González et al., 2011; Vaithilingam et al., 2012)**.

A travers le mécanisme de phagocytose des cellules apoptotiques dépendant de C1q que j'ai décrit au cours de cette introduction, nous pouvons nous demander si l'amibe ne se comporterait pas comme un macrophage, et utiliserait le même type de mécanisme en s'appropriant les ressources de l'hôte (C1q) pour favoriser la phagocytose de nos propres cellules. La CRT d'*Entamoeba* interagissant uniquement avec des cellules rendues apoptotiques et non avec des cellules vivantes, elle pourrait participer à l'interaction soit de manière directe avec la PS ou la CRT à la surface des cellules apoptotiques, soit *via* C1q en suivant le modèle proposé chez les mammifères.

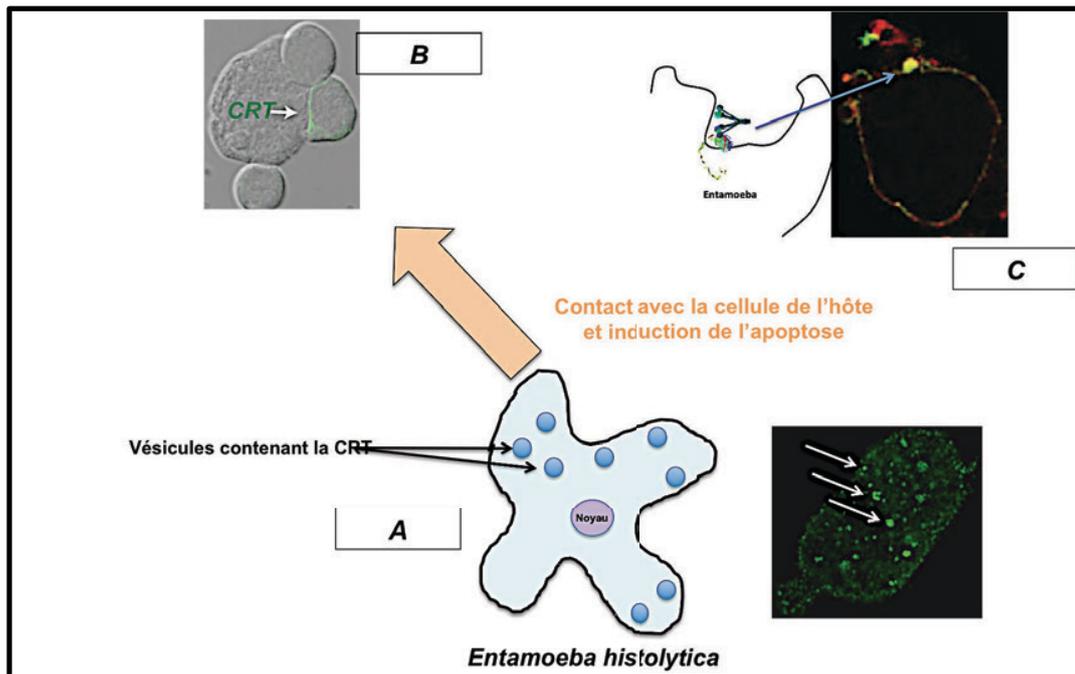


FIGURE 1.V.4 : PARTICIPATION DE LA CRT A LA PHAGOCYTOSE AMEBIQUE (XIMENEZ ET AL., 2014)

(A) La calréticuline exerce sa fonction chaperonne dans des vésicules proches de la membrane plasmique de l'amibe, qui jouent le rôle de réticulum. (B) Lors d'un contact avec une cellule de l'hôte, l'amibe induit sa mort par apoptose et l'avale par phagocytose. Durant l'étape initiale de phagocytose, la CRT est exportée à la surface de l'amibe au niveau de la cupule phagocytaire où elle interagit avec C1q (C).

Projet de thèse

Le sujet de ma thèse a été principalement focalisé sur l'étude des calréticulines de deux parasites (*T. cruzi* et *E. histolytica*) et de leur interaction avec la molécule C1q. J'ai pu vous présenter dans cette introduction le contexte de l'étude, montrant entre autres la variété des fonctions de C1q et des calréticulines, les données de structure précédemment obtenues sur les domaines globulaires de C1q et des calréticulines de mammifère. J'ai pu situer le rôle des CRTs et de C1q dans le contexte plus global des pathologies liées à ces deux parasites. Finalement, comme l'illustre la figure I.VI.I, l'étude de ces deux modèles pourrait éventuellement nous permettre de décomposer les deux éléments fonctionnels de l'interaction entre C1q et la calréticuline humaine dans le cadre de l'efférocytose, en suivant l'hypothèse que l'interaction C1q/CRT à la surface des cellules apoptotiques est mimée par l'interaction entre C1q et CRT à la surface de *T. cruzi* et que l'interaction entre la CRT et C1q au niveau de la cupule phagocytaire de l'amibe ressemble à l'interaction C1q/CRT à la surface des macrophages..

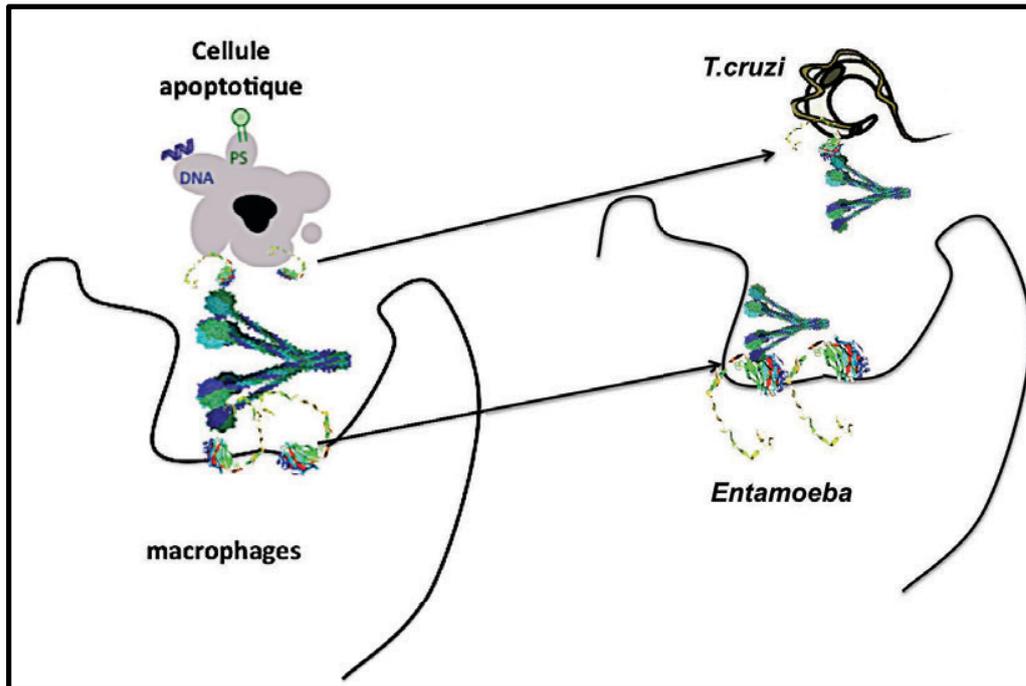


FIGURE 1.VI.1: ILLUSTRATION DU PROJET DE THESE

Décomposition des fonctions portées par la calréticuline humaine à travers deux modèles d'études, *T. cruzi* jouant le rôle de la cellule apoptotique et *E. histolytica* le rôle du macrophage.

Dans ce cadre, mes objectifs principaux ont été de:

- a) Définir plus précisément les domaines impliqués dans l'interaction entre C1q et CRT des parasites et comparer ces interactions.
- b) Obtenir des informations structurales sur les domaines des CRTs interagissant avec C1q, si possible à haute résolution.
- c) Comparer finement les structures/séquences des CRT humaine et de parasites, qui sont très éloignés sur le plan phylogénétique, pour identifier des similarités en lien avec leur rôle commun de chaperonne et leur interaction avec C1q.
- d) Produire en grande quantité et caractériser une forme recombinante des régions globulaires de C1q (GRC1q), ce qui ouvrirait de nouvelles perspectives en particulier pour les analyses structurales de complexes de cette protéine.

2. Matériels et Méthodes

I. Matériel génétique et milieux de cultures

Les différentes souches cellulaires, milieux de cultures et vecteurs plasmidiques utilisés au cours de ma thèse pour les clonages et productions des différentes protéines recombinantes sont décrits dans les **tableaux 2.I.1 et 2.I.2**, et les **figures 2.I.1 et 2.I.2**.

| | | |
|----------------------------|--|---|
| SOUCHES CELLULAIRES | <u><i>E. coli</i> DH5α</u> (Invitrogen) | <i>F</i> - Φ 80 <i>lacZ</i> AM15 Δ (<i>lacZYA-argF</i>) U169 <i>recA1 endA1 hsdR17</i> (<i>rK</i> -, <i>mK</i> +) <i>phoA supE44</i> λ - <i>thi-1 gyrA96 relA1</i> . Cette souche a été utilisée pour amplifier et obtenir en grande quantité les ADN plasmidiques. |
| | <u><i>E. coli</i> XL10 GOLD</u> (Agilent technologies) | <i>TetrD (mcrA)183 D(mcrCB-hsdSMR-mrr)173 endA1 supE44 thi-1 recA1 gyrA96 relA1 lac Hte [F' proAB lacIqZDM15 Tn10 (Tetr) Amy Camr. (Jerpseth et al., 1993 & 1997), (Bullock WO et al., 1987).</i> Cette souche ultra compétente dérivée de la souche XL2-Blue-MRF (Agilent Technologies) a été utilisée pour amplifier les constructions de plasmides obtenus par PCR pour générer les différents fragments de calréticuline et les GRC1q recombinantes. |
| | <u><i>E. coli</i> BL21 (DE3)</u> | <i>huA2 [lon] ompT gal (λ DE3) [dcm] ΔhsdS λ DE3 = λ sBamHIo ΔEcoRI-B int::(<i>lacI::PlacUV5::T7 gene1</i>) i21 Δnin5.</i> Elle possède une copie chromosomique du gène codant pour la T7 ARN polymérase, sous le contrôle du promoteur LacUV5 naturellement réprimé par la protéine LacI dont le gène est également présent dans le génome de cette souche, et inductible par l'IPTG. Elle est utilisée pour l'expression de protéines dont la séquence codante est placée sous le contrôle du promoteur T7. |
| | <u>HEK 293-F</u> (Life Technologies) | Les cellules HEK 293 Freestyle sont des cellules embryonnaires provenant de l'épithélium de rein humain dérivées de la lignée 293. Elles ont été utilisées pour produire la protéine recombinante GRC1q. |

TABLEAU 2.I.1 : TABLEAU PRESENTANT LES DIFFERENTES SOUCHES CELLULAIRES UTILISEES POUR LE CLONAGE ET L'EXPRESSION DES FRAGMENTS DE CALRETICULINE ET DES GRC1Q RECOMBINANTES

| | |
|---|---|
| <p style="text-align: center;">Milieu riche L B (Luria Bertani)</p> | <p>Le LB est constitué de tryptone (20 g/L), extrait de levure (5 g/L), NaCl (5 g/L). Le pH est ajusté à 7,5. Pour le milieu solide, de l'agar est ajouté (20 g/L) avant stérilisation. Les milieux commerciaux LB Lennox Agar (BD Biosciences) et LB Lennox Broth Base (Invitrogen) ont également été utilisés. Ces milieux sont utilisés pour la culture bactérienne et la production de protéines recombinantes en grande quantité, ainsi que pour la transformation des cellules <i>E. coli</i>. L'ensemble de ces milieux est autoclavé pendant 30 minutes à 120°C sous une pression de 1bar.</p> |
| <p style="text-align: center;">Milieu riche NZY⁺</p> | <p>Le milieu NZY⁺ est composé de NZ amine (hydrolysate de caséine) (10 g/L), d'extrait de levure (5 g/L), NaCl (5 g/L) et le pH est ajusté à 7,5 par ajout de NaOH. Après stérilisation par autoclave 30 minutes à 120°C sous une pression de 1 bar, le milieu est supplémenté avec 12,5 mM (final) de MgCl₂, 12,5 mM (final) de MgSO₄ et du glucose 20 mM (final). Ce milieu est utilisé pour l'étape de régénération lors de la transformation des souches de <i>E. coli</i> avec les différents ADN plasmidiques.</p> |
| <p style="text-align: center;">Milieu minimum enrichi M9</p> | <p>Le milieu M9 enrichi est un milieu dont la composition chimique est contrôlée, utilisé pour la production de la calréticuline de <i>T. cruzi</i> marquée ¹⁵N et ¹³C pour les expériences de RMN. Il est composé de Na₂HPO₄.7H₂O (10 g/L), KH₂PO₄ (3 g/L), NaCl (0,5 g/L) et NH₄Cl (ou ¹⁵NH₄Cl selon le marquage choisi) (1 g/L), à pH compris entre 6,8 et 7,2 et autoclavé pendant 30 minutes à 120°C sous une pression de 1 bar. Le milieu contient également une source de carbone avec 2 mM de glucose (ou ¹³glucose selon le marquage souhaité), ainsi que de MgSO₄ (1 mM), CaCl₂ (100 µM), MnCl₂ (100 µM), des métaux : ZnSO₄ (50 µM), FeCl₃ (50 µM) et un cocktail de vitamines composé de pyridoxine (1 mg/L), biotine (1 mg/L), panthothénate héli calcium (1 mg/L), acide folique (1 mg/L), chlorure de choline (1 mg/L), niacinamide (1 mg/L), riboflavine (100 µg/L), thiamine (5 mg/L).</p> |
| <p style="text-align: center;">Milieu Freestyle 293 (Life Sciences Technologies)</p> | <p>Le milieu GIBCO™ FreeStyle™ 293 est un milieu sans sérum utilisé pour la culture des cellules de mammifères HEK 293-F pour la production de la protéine recombinante GRC1q.</p> |

TABEAU 2.I.2 : TABLEAU PRESENTANT LES DIFFERENTS MILIEUX DE CULTURES UTILISES POUR LE CLONAGE ET L'EXPRESSION DES FRAGMENTS DE CALRETICULINE ET DES GRC1Q RECOMBINANTES

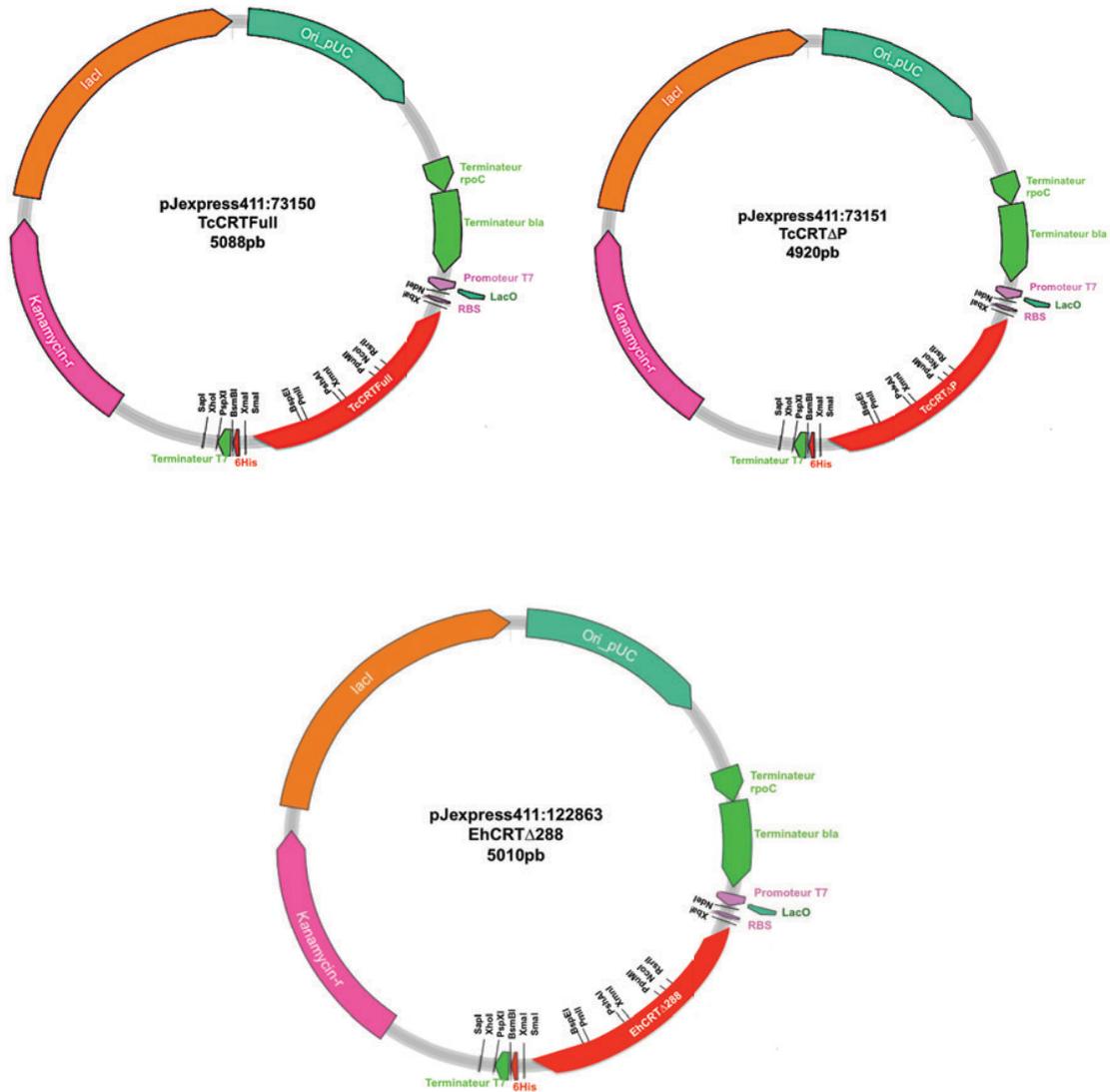


FIGURE 2.I.1 : CARTE VECTORIELLE DES PLASMIDES D'EXPRESSION BACTERIENS UTILISES

Les plasmides pJexpress411 :122863-EHCRTΔ288, pJexpress411 :73150-TcCRTFull et pJexpress411 :73151-TcCRTΔP proviennent de la société DNA2.0. Ils possèdent un gène de résistance à la kanamycine (kanamycin-r) permettant la sélection des bactéries transformées, ainsi qu'une origine de répllication (Ori_pUC) permettant la répllication du plasmide dans *E. coli*. Les deux sites de restriction utilisés par DNA2.0 pour introduire les gènes codant pour les fragments de CRTs (XbaI et XhoI) sont présents dans une cassette de clonage présentant plusieurs sites de restriction uniques et sont situés entre les promoteurs et terminateurs reconnus par l'ARN polymérase du bactériophage T7 et en amont d'une séquence codant pour une étiquette de 6 histidines. Une séquence RBS permet la fixation du ribosome et la traduction de l'ARN en protéine. Les tailles totales de ces plasmides sont de 5010, 5088 et 4920 paires de bases respectivement. Les différents fragments de EhCRT et TcCRT ont été obtenus par PCR à partir de cette construction.

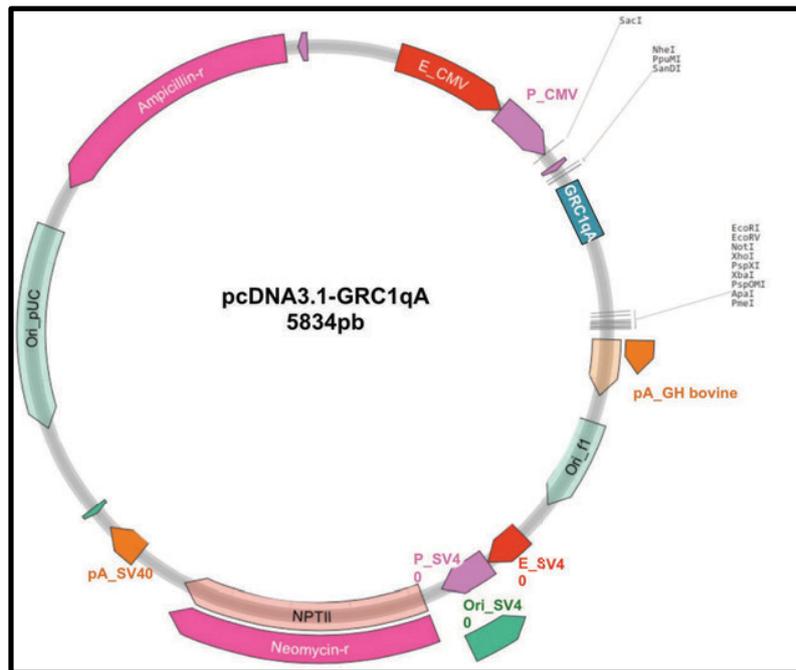


FIGURE 2.I.2 : CARTE VECTORIELLE DU PLASMIDE D'EXPRESSION EN CELLULES DE MAMMIFERE UTILISE

Le plasmide pcDNA3.1(-) (Invitrogen) contenant GRC1qA est un vecteur navette utilisé pour le clonage et l'expression en cellules de mammifères HEK-293F de la forme recombinante de la protéine GRC1q. Il contient des sites de restrictions uniques présents dans une cassette de clonage permettant l'insertion d'un gène d'intérêt, ainsi que les éléments essentiels à sa réplication chez *E. coli* : une origine de réplication pour *E. coli* (Ori_pUC), pour le phage M13 (Ori_f1), un gène de résistance à l'ampicilline permettant la sélection des bactéries transformées. Pour l'expression en cellules de mammifères le gène codant pour les GRC1q recombinantes est introduit en aval d'un promoteur fort (P_CMV le promoteur du cytomégalovirus reconnu par l'ARN polymérase III des cellules de mammifères) et en amont d'un signal de polyadénylation de l'hormone de croissance bovine (pA_GH bovine). Un gène de résistance à la neomycine est présent (Neomycin-r) permettant d'exercer une pression de sélection sur les cellules afin d'intégrer le plasmide à leur ADN génomique pour obtenir des lignées stables. Le plasmide peut également être répliqué de manière épisomale sans être intégré au génome grâce à la machinerie de réplication et transcriptionnelle du virus Simien (E_SV40, P_SV40, pA_SV40 et Ori_SV40).

Ce plasmide possède également la séquence codant pour la chaîne A des régions globulaires de C1q précédées du peptide signal de la chaîne A, clonées entre les sites de restriction *NdeI* et *EcoRI* (clonage réalisé au laboratoire par Isabelle Bally).

II. Méthodes de biologie moléculaire

II.1. Amplification d'une séquence d'intérêt par PCR

II.1.a. Principe

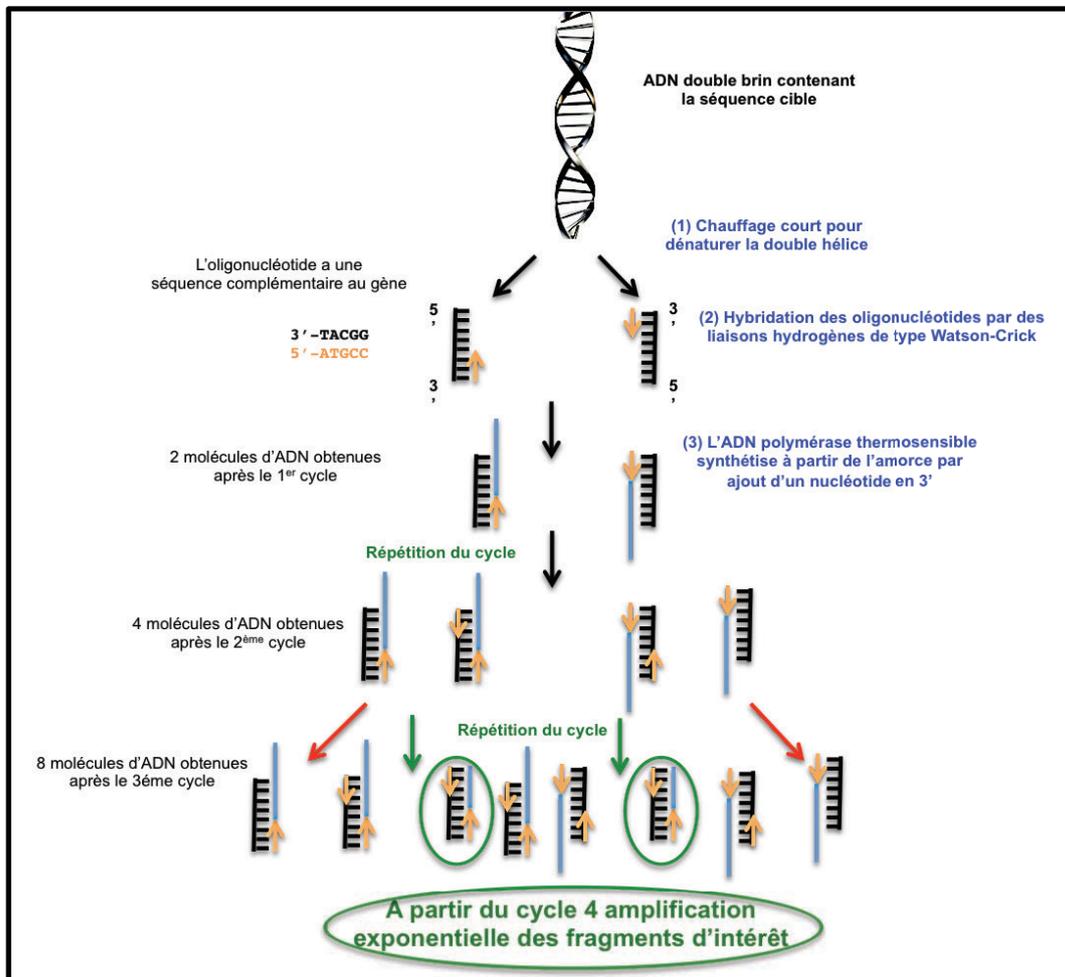


FIGURE 2.II.1 : PRINCIPE DE LA TECHNIQUE PCR (POLYMERASE CHAIN REACTION)

La Polymerase Chain Reaction (PCR), développée par K. Mullis en 1985, permet de répliquer en l'amplifiant un fragment d'ADN à partir d'un ADN matrice. La spécificité de la réaction est assurée par deux oligonucléotides s'appariant spécifiquement aux extrémités 3' du gène sur chacun des deux brins de l'ADN matrice. Une première étape de chauffage à 95°C suivie d'une baisse de la température permet la dénaturation de la double hélice d'ADN, puis l'hybridation des oligonucléotides (**1 et 2**) sur la matrice. Une augmentation de la température à 72°C, température optimale d'activité de la polymérase thermostable Vent, permet l'élongation des brins sur le principe de la répllication, les nucléotides servant d'amorce (**3**). Ce cycle composé de trois étapes est répété, et à partir du 4^{ème} cycle le fragment d'ADN d'intérêt est obtenu et sera amplifié de manière exponentielle.

II.1.b. Protocole et oligonucléotides utilisés pour l'amplification de la séquence codant pour les chaînes B et C des GRC1q dans le plasmide d'expression en cellule de mammifère

L'expression des GRC1q sous forme recombinante monocaténaire avait été entreprise au laboratoire ces dernières années sur la base de la stratégie décrite pour l'adiponectine (Ge et al., 2010; Min et al., 2012). De ce fait, nous disposons au laboratoire d'un gène synthétique, comportant les 3 chaînes reliées 2 à 2 par des linkers dans l'ordre A-C-B dans le vecteur pNT-Bac utilisé pour l'expression en cellules d'insectes. En parallèle, les trois chaînes avaient été clonées séparément dans le plasmide pcDNA3.1.

La séquence codant pour un fragment linker A-C-gC1qC-linker C-B-gC1qB est amplifiée par PCR à partir du plasmide pNT-Bac-GRC1q selon le principe présenté dans la section précédente. L'utilisation de ce fragment comme méga-primer pour insérer la séquence codant pour les chaînes B et C des GRC1q dans le plasmide pcDNA3.1 en aval de la séquence codant la chaîne A par une méthode de mutagenèse dirigée (dont je décrirai le principe dans la section suivante) a nécessité l'ajout d'une extrémité 5' débordante à chacun des oligonucléotides. Ces dernières sont complémentaires de la séquence codant pour la chaîne A et de la séquence du plasmide pcDNA3.1 située en aval de la chaîne A, respectivement (Tableau 2.II.7 et Figure 2.II.2). Le mélange réactionnel de PCR utilisé et le programme PCR sont présentés respectivement dans les tableaux 2.II.1 et 2.II.2.

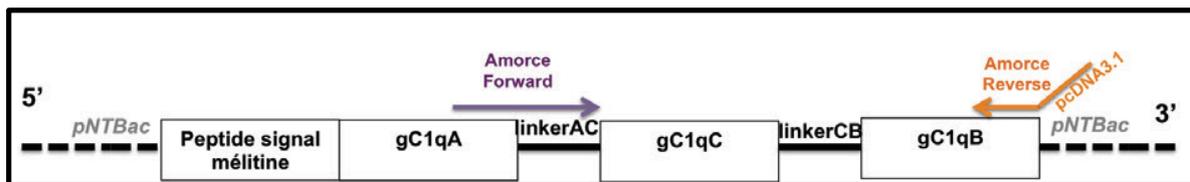


FIGURE 2.II.2 : SCHEMA DECRIVANT LA STRATEGIE D'AMPLIFICATION DU GENE CODANT POUR GRC1Q_BC AFIN D'OBTENIR UN MEGA-PRIMER

Le choix des oligonucléotides permet lors de l'amplification du gène codant pour les chaînes B et C des GRC1q d'ajouter des extrémités débordantes complémentaires du gène codant pour la chaîne A déjà dans le plasmide (violet) et de la séquence du plasmide placée en aval de la chaîne A (orange).

| | Volume | Concentration finale |
|--------------------------------------|--------|----------------------|
| Oligo Forward | 1 µL | 1 µM |
| Oligo Reverse | 1 µL | 1 µM |
| dNTP | 1,6 µL | 0,4 mM |
| pNTBac_GRC1q (GenScript) | 4 µL | 2 ng/µL |
| MgSO₄²⁻ | 2 µL | 2 mM |
| Vent Polymérase | 1 µL | 0,02 U/µL |
| Tampon Vent Polymérase 10x | 10 µL | 1x |
| H₂O qsp 100 µL | 77 µL | - |

TABLEAU 2.II.1 : MELANGE DE PCR POUR L'AMPLIFICATION DE LA SEQUENCE CODANT GRC1Q_BC

| Phases | Température (°C) | Temps (minutes) | Nombre de cycles |
|---|------------------|-----------------|------------------|
| Dénaturation | 95 | 10 | 1 |
| AJOUT DE LA VENT POLYMERASE ET dNTPs | | | |
| Dénaturation initiale | 95 | 2 | 1 |
| Dénaturation | 95 | 0,5 | 30 |
| Hybridation | 63 | 0,5 | |
| Elongation | 72 | 1 | |
| Elongation finale | 72 | 5 | 1 |

TABLEAU 2.II.2 : PROGRAMME DE PCR POUR L'AMPLIFICATION DE LA SEQUENCE CODANT GRC1Q_BC

II.1.c. Purification de la séquence amplifiée

Après migration sur gel d'agarose 1%, la bande correspondant au fragment d'intérêt amplifié est révélée sous UV en présence de SYBR Green[®]. Le gel est découpé et le fragment d'ADN est extrait du gel à l'aide du kit « Qiagen Gel Extraction Kit » de Qiagen selon les indications du fournisseur. Cette étape permet d'éliminer les nucléotides, les enzymes, sels, le SYBR Green[®], l'agarose et toute les impuretés présentes dans l'échantillon. L'ADN obtenu est conservé à -20°C.

II.2. Délétion/Insertion de séquence d'intérêt par PCR

II.2.a. Principe

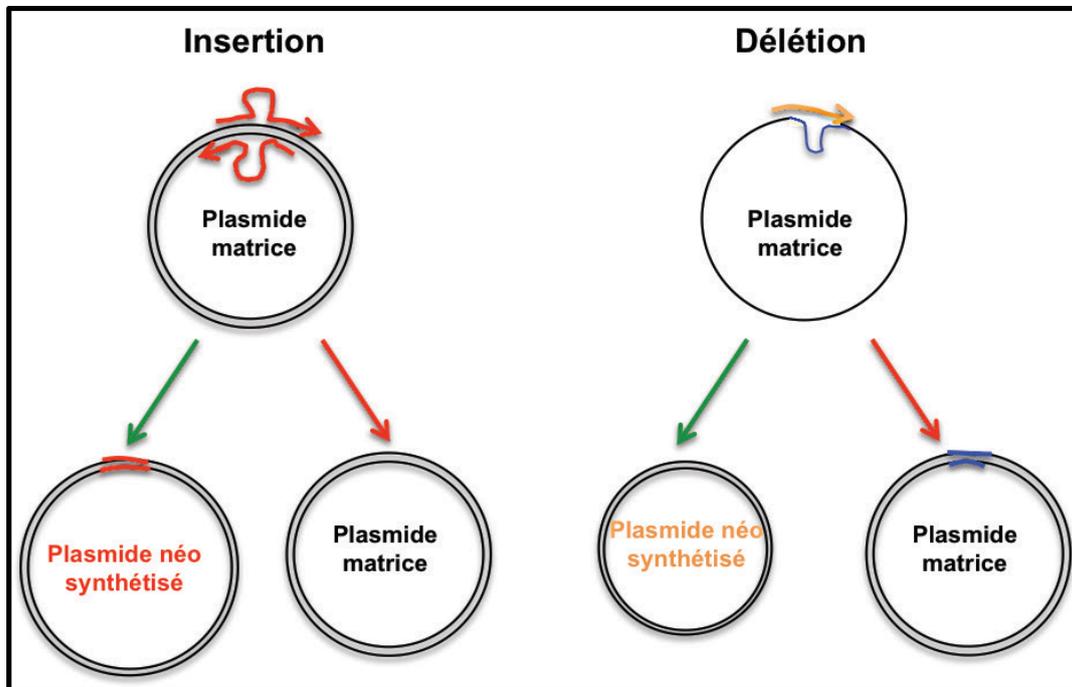


FIGURE 2.II.3 : PRINCIPE DE L'INSERTION/DELETION D'UNE SEQUENCE PAR PCR (MUTAGENESE DIRIGEE)

L'insertion/délétion d'une séquence d'intérêt repose sur la technique de la mutagenèse dirigée, dérivée de la technique de PCR. Comme pour la PCR classique le plasmide matrice à amplifier est mis en présence de deux amorces d'ADN complémentaires, contenant soit un segment d'ADN supplémentaire non apparié à l'ADN matrice (**insertion**), soit une délétion volontaire d'une portion de la séquence portée par l'amorce vis à vis de la séquence complémentaire à appairer sur l'ADN matrice (**délétion**). A nouveau une étape de chauffage à 95°C suivie d'une baisse de la température permet la dénaturation de la double hélice d'ADN, puis l'hybridation des oligonucléotides sur la matrice. Une augmentation de la température à 68°C, température optimale d'activité de la Pfu polymérase thermostable, permet l'élongation des brins sur le principe de la réplication, les nucléotides servant d'amorce.

Les différentes expériences de mutagenèse dirigée décrites ci-après permettant l'insertion/délétion de séquences d'ADN d'intérêt ont été réalisées avec le kit QuickChange XL Site-Directed Mutagenesis (Agilent Technologies). Selon le cas de figure, des optimisations ont été apportées au protocole standard du fournisseur.

II.2.b. Protocole utilisé pour l'insertion de la séquence codant pour les chaînes B et C des GRC1q dans le plasmide pcDNA3.1-GRC1q_A

La stratégie mise en place pour cloner la séquence codant le fragment linkerAC-gC1qC-linkerCB-gC1qB dans le plasmide pcDNA3.1 dans lequel se trouve déjà la séquence codant pour la chaîne A des GRC1q précédée par le peptide signal de la chaîne A repose sur un protocole en deux étapes permettant l'insertion de grands fragments d'ADN (utilisés en tant que méga-primer) dans un plasmide sans passer par une étape de clonage ((Geiser et al., 2001), thèse de Mickael Jacquet, 2012). Le produit d'amplification PCR obtenu est donc utilisé comme amorce et mélangé avec la matrice avant de subir une élongation et une amplification par PCR (Figure 2.II.4). Le mélange réactionnel de PCR utilisé et le programme PCR sont présentés respectivement dans les tableaux 2.II.3 et 2.II.4.

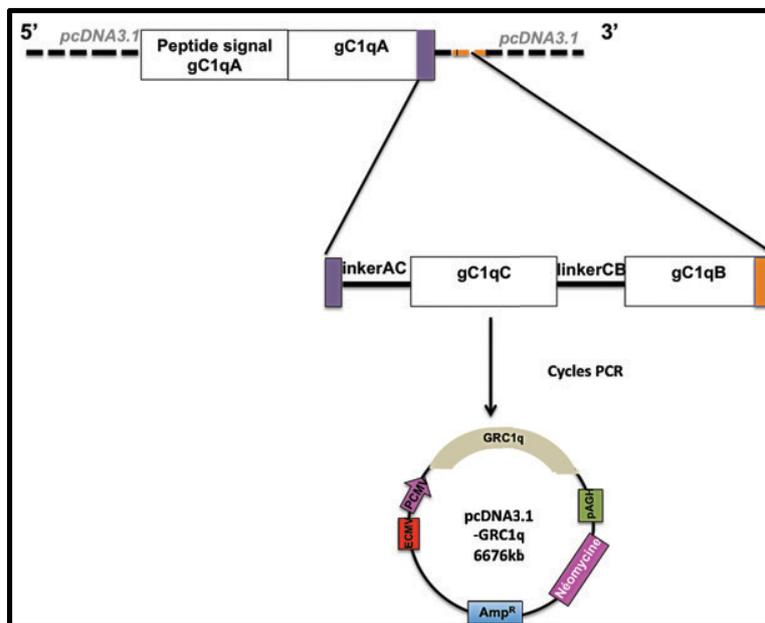


FIGURE 2.II.4 : SCHEMA PRESENTANT LE PROTOCOLE D'INSERTION DU GENE CODANT LES CHAINES A ET B GRC1Q DANS LE PLASMIDE PCDNA3.1-GRC1Q-A

Ce fragment utilisé comme méga primer sert d'amorce pour répliquer et amplifier le plasmide à travers différents cycles PCR. Ces cycles PCR permettent l'intégration de la séquence d'intérêt au plasmide. Les extrémités débordantes ajoutées lors de l'étape de PCR s'apparient avec la séquence codant pour gC1qA (violet) et la séquence du plasmide pcDNA3.1 (orange).

| | Volume | Concentration finale |
|--|---------|----------------------|
| pcDNA3.1-GRC1qA (50 ng/μL) | 1 μL | 1 ng/μL |
| dNTPmix | 1,5 μL | 0,4 mM |
| Mégaprimer (produit PCR) (16 ng/μL) | 12,5 μL | 4 ng/μL |
| Pfu Polymérase | 1 μL | 0,05 U/μL |
| QuickSolution | 3 μL | 6% |
| Tampon QuickChange 10x | 5 μL | 1x |
| H₂O qsp 50 μL | 26 μL | - |

TABLEAU 2.II.3 : MELANGE DE PCR POUR L'AMPLIFICATION DE LA SEQUENCE CODANT GRC1Q_BC

| Phases | Température (°C) | Temps (minutes) | Nombre de cycles |
|--|------------------|-----------------|------------------|
| Dénaturation | 95 | 10 | 1 |
| AJOUT DE LA Pfu POLYMERASE ET dNTPs | | | |
| Dénaturation initiale | 95 | 0,5 | 1 |
| Dénaturation | 95 | 0,5 | 5 |
| Hybridation | 52 | 1 | |
| Elongation | 68 | 10 | |
| Dénaturation initiale | 95 | 0,5 | 13 |
| Dénaturation | 95 | 0,5 | |
| Hybridation | 55 | 1 | |
| Elongation | 68 | 10 | |
| Elongation finale | 68 | 7 | 1 |

TABLEAU 2.II.4: PROGRAMME DE PCR POUR L'AMPLIFICATION DE LA SEQUENCE CODANT GRC1Q_BC

II.2.c. Protocole utilisé pour les délétions introduites dans les calréticulines

Comme pour l'insertion du gène codant les GRC1q monocaténaire, les délétions introduites dans les différents fragments de calréticuline reposent sur une optimisation du protocole de base du kit QuickChange (Qiagen). Les grandes délétions nécessitent des oligonucléotides de grande taille, l'optimisation se situe donc au niveau de la phase d'amplification par PCR divisée en deux étapes (Wang and Malcolm, 1999) et thèse de Mickael Jacquet 2012, pour optimiser l'efficacité de la réaction et limiter l'oligomérisation des deux amorces entre elles. Deux réactions de PCR indépendantes sont d'abord effectuées avec un seul oligonucléotide par réaction. Cette stratégie permet de répliquer chaque brin du plasmide indépendamment et de générer de longs fragments d'ADN qui seront utilisés comme mégaprimers lors de la troisième PCR. Le choix des amorces, ainsi que le mélange et le programme PCR utilisés sont décrits Figure 2.II.5, et tableaux 2.II.5 et 2.II.6.

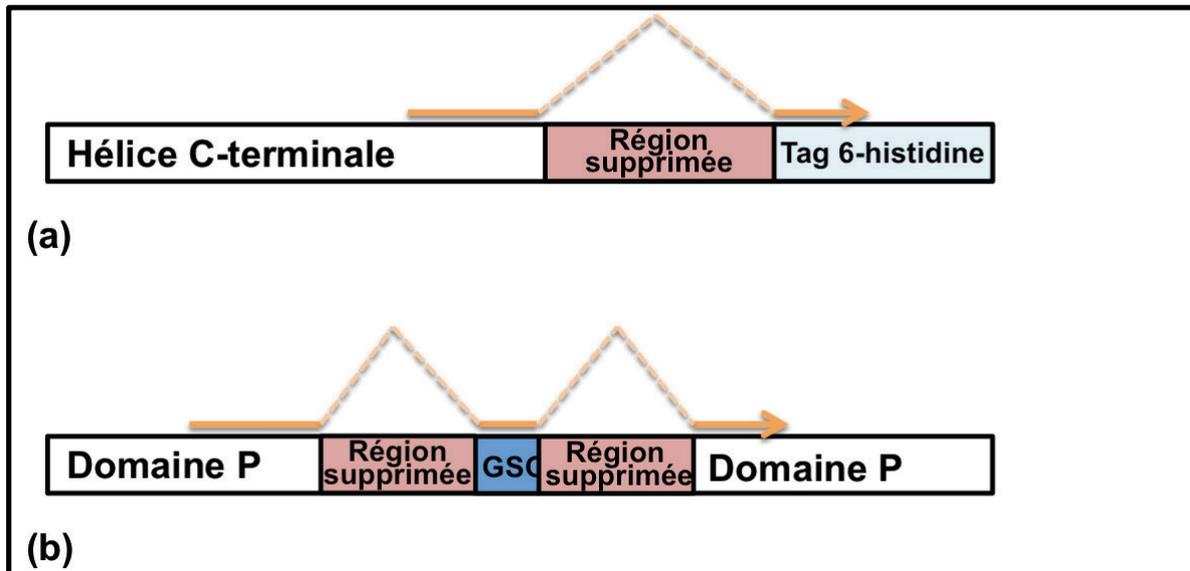


FIGURE 2.II.5 : SCHEMA PRESENTANT LA STRATEGIE UTILISEE POUR LE CHOIX DES OLIGONUCLEOTIDES POUR SUPPRIMER DES REGIONS DES DIFFERENTES CALRETICULINES PAR PCR.

Pour les différentes troncations de l'hélice C-terminale des calréticulines de *T. cruzi* et *E. histolytica*, les oligonucléotides ont été choisis pour s'hybrider sur la séquence codant pour l'hélice en amont de la séquence à supprimer et systématiquement sur la séquence codant pour le tag histidine situé en aval (a). Pour les troncations du domaine P, les oligonucléotides ont été choisis dans le but de réaliser deux délétions en une seule réaction de PCR en amont et en aval du linker GSG. Pour cela la séquence de l'oligonucléotide s'apparie avec la séquence codant pour le domaine P en amont de la première délétion, avec la séquence codant le linker GSG (située au centre de l'amorce) et enfin avec la séquence codant pour le domaine P en aval de la seconde délétion (b).

| | Volume tube 1 | Volume tube 2 | Concentration finale (par tube) |
|---|---------------|---------------|---------------------------------|
| Plasmide matrice (25 ng/μL) | 1 μL | 1 μL | 1 ng/μL |
| dNTPmix | 1 μL | 1 μL | 0,3 mM |
| Amorce sens (125 ng/μL) | - | 1 μL | 5 ng/μL |
| Amorce anti-sens (125 ng/μL) | 1 μL | - | 5 ng/μL |
| Pfu Polymérase | 0,5 μL | 0,5 μL | 0,05 U/μL |
| QuickSolution | 1,5 μL | 1,5 μL | 3% |
| QuickChange Buffer 10x | 2,5 μL | 2,5 μL | 1x |
| H₂O qsp 25 μL | 17,5 μL | 17,5 μL | - |

TABLEAU 2.II.5 : MELANGE DE PCR POUR LES DELETIONS INTRODUITES DANS LES CALRETICULINES

| Phases | Température (°C) | Temps (minutes) | Nombre de cycles |
|---|------------------|-----------------|------------------|
| Dénaturation | 95 | 10 | 1 |
| AJOUT DE LA Pfu POLYMERASE ET dNTPs | | | |
| Dénaturation initiale | 95 | 0,5 | 1 |
| Dénaturation | 95 | 0,5 | 5 |
| Hybridation | 55 | 1 | |
| Elongation | 68 | 9 | |
| Elongation finale | 68 | 8 | 1 |
| MELANGE TUBE 1 + TUBE 2 + 0,25 μL Pfu Polymérase | | | |
| Dénaturation initiale | 95 | 1 | 18 |
| Dénaturation | 95 | 1 | |
| Hybridation | 60 | 1 | |
| Elongation | 68 | 7 | |
| Elongation finale | 68 | 7 | 1 |

TABLEAU 2.II.6 : PROGRAMME DE PCR POUR LES DELETIONS INTRODUITES DANS LES CALRETICULINES

II.2.d Sélection des ADN plasmidiques néosynthétisés

Comme décrit dans la **figure 2.II.3**, à la fin d'une réaction de mutagenèse dirigée, un mélange de deux types d'ADN plasmidiques cohabitent : l'ADN matrice et l'ADN néosynthétisé contenant les modifications souhaitées. L'ADN matrice possédant également un gène de résistance à un antibiotique, utiliser le mélange réactionnel tel quel pour transformer les bactéries risque d'entraîner l'apparition de faux positifs. Il est alors important de se débarrasser de l'ADN matrice. La méthode repose sur les modifications apportées à l'ADN matrice par les cellules dans lesquelles il a été répliqué. En effet *E. coli* méthyle l'ADN, alors que l'ADN néosynthétisé *in vitro* ne possède aucune méthylation. Le protocole consiste à ajouter 1 µL de l'enzyme DpnI fournie avec le kit (une endonucléase reconnaissant spécifiquement les séquences GATC méthylées) et à incuber 1h à 37°C. Le protocole est répété deux fois.

II.3. Oligonucléotides

Les différentes amorces utilisées pour la construction des différents vecteurs d'expression décrits précédemment ont été synthétisées par Eurogentec S.A. Leur température de fusion (T_m) a été calculée à l'aide de la formule fournie par Agilent Technologies et optimisée pour le kit QuickChange XL Site-Directed Mutagenesis ($T_m = 81.5 + 0.41(\%GC) - (675/N)$, avec N le nombre de bases contenues dans l'amorce sans tenir compte des bases non appariées). La liste des amorces est répertoriée dans les **tableaux 2.II.7, 2.II.8, 2.II.9**.

| Noms | Séquences | Description | T_m (°C) |
|-----------------------|--|--|------------|
| GRC_GRBForward | GGC TTC CTC ATC TTC CCA TCT GCC GGC AGT GGC AAG CAG | Amorce sens utilisée pour amplifier la séquence codant pour les chaînes B et C des GRC1q en ajoutant une extrémité débordante s'appariant sur la séquence codant la chaîne A et le linker A-C | 75 |
| GRC_GRBReverse | CCA CTG TGC TGG ATA TCT GCA GAA TTC TCA GGC CTC CAT ATC TGG AAA GAG CAG G | Amorce antisens utilisée pour amplifier la séquence codant pour les chaînes B et C des GRC1q en ajoutant une extrémité débordante s'appariant sur la séquence du plasmide pcDNA3.1 | 74 |

TABLEAU 2.II.7 : OLIGONUCLEOTIDES UTILISES POUR L'AMPLIFICATION DE LA SEQUENCE CODANT POUR LES CHAINES B ET C DES GRC1Q (LA SEQUENCE EST DONNEE DE 5' EN 3')

| Noms | Séquences | Description | Tm (°C) |
|--------------------|---|--|---------|
| TcCRTΔPA301Forward | GAA CGT GAG AAG GGC GAC CAC CAT CAC C | Suppression hélice Cter entre le résidu 301 et le tag GDHHHHHH. S'hybride sur la séquence en amont du résidu 301 et sur le tag | 82 |
| TcCRTΔPA301Reverse | G GTG ATG GTG GTC GCC CTT CTC ACG TTC | | |
| TcCRTΔPA290Forward | G GAT ATG GAA AAA GAG AAA GGC GAC CAC CAT CAC C | Suppression hélice Cter entre le résidu 290 et le tag GDHHHHHH. S'hybride sur la séquence en amont du résidu 290 et sur le tag | 82 |
| TcCRTΔPA290Reverse | G GTG ATG GTG GTC GCC TTT CTC TTT TT CCA TAT CC | | |
| TcCRTΔPA288Forward | G GAG GAT ATG GAA AAA GGC GAC CAC CAT CAC C | Suppression hélice Cter entre le résidu 288 et le tag GDHHHHHH. S'hybride sur la séquence en amont du résidu 288 et sur le tag | 83 |
| TcCRTΔPA288Reverse | G GTG ATG GTG GTC GCC TTT TTC CAT ATC CTC C | | |
| TcCRTΔPA281Forward | GCG GAA GCG GAC GCC CTG AAA GGC GAC CAC CAT CAC CAT C | Suppression hélice Cter entre le résidu 281 et le tag GDHHHHHH. S'hybride sur la séquence en amont du résidu 281 et sur le tag | 91 |
| TcCRTΔPA281Reverse | GAT GGT GAT GGT GGT CGC CTT TCA GGG CGT CCG CTT CGG C | | |
| TcCRTΔPA278Forward | GCG CTG AAG AAG GCG GAA GCG GGC GAC CAC CAT CAC CAT C | Suppression hélice Cter entre le résidu 278 et le tag GDHHHHHH. S'hybride sur la séquence en amont du résidu 278 et sur le tag | 91 |
| TcCRTΔPA278Reverse | GAT GGT GAT GGT GGT CGC CCG CTT CCG CCT TCT TCA GGC C | | |
| TcCRTΔPA276Forward | CGG TGG CCT GAA GAA GGC GGG CGA CCA CCA TCA CC | Suppression hélice Cter entre le résidu 276 et le tag GDHHHHHH. S'hybride sur la séquence en amont du résidu 276 et sur le tag | 90 |
| TcCRTΔPA276Reverse | GGT GAT GGT GGT CGC CCG CCT TCT TCA GGC CAC CG | | |
| TcCRTΔPA274Forward | GGT GGC CTG AAG AGG GGC GAC CAC CAT CAC C | Suppression hélice Cter entre le résidu 274 et le tag GDHHHHHH. S'hybride sur la séquence en amont du résidu 274 et sur le tag | 87 |
| TcCRTΔPA274Reverse | G GTG ATG GTG GTC GCC CTT CTT CAG GCC ACC | | |
| TcCRTΔPA270Forward | GGT GGA GAA AAC CTA CGG TGG CGA CCA CC | Suppression hélice Cter entre le résidu 270 et le tag GDHHHHHH. S'hybride sur la séquence en amont du résidu 270 et sur le tag | 80 |
| TcCRTΔPA270Reverse | GGT GGT CGC CAC CGT AGG TTT TCT CCA CC | | |
| TcCRTΔPA268Forward | CTG GTG GAG AAA ACC GGC GAC CAC CAT CAC | Suppression hélice Cter entre le résidu 268 et le tag GDHHHHHH. S'hybride sur la séquence en amont du résidu 268 et sur le tag | 80 |
| TcCRTΔPA268Reverse | GTG ATG GTG GTC GCC GGT TTT CTC CAC CAG | | |

TABEAU 2.II.8 : OLIGONUCLEOTIDES UTILISES POUR INTRODUIRE DES DELETIONS DANS L'HELICE C-TERMINALE DE LA CALRETICULINE DE *T. CRUZI* (LA SEQUENCE EST DONNEE DE 5' EN 3')

| Noms | Séquences | Description | Tm (°C) |
|------------------------------|--|---|-----------|
| TcCRTΔPbisΔ288Forward | AGC CTG TTG CTG CCG CGT GGC AGC GGT TTT GTG GAG GAT TCT GAA | Suppression de deux régions du domaine P en amont (arginine 182) et en aval (phénylalanine 220) du linker GSG. S'hybride sur la séquence en amont du résidu 182, le linker, et la séquence en aval du résidu 280 | 90 |
| TcCRTΔPbisΔ288Reverse | TTC AGA ATC CTC CAC AAA ACC GCT GCC ACG CGG CAG CAA CAG GCT | | |
| TcCRTΔPterΔ288Forward | CCG CGT GAG ATT GTG GAT GGC AGC GGT ATC CCG AAC CCG GAT TTT | Suppression de deux régions du domaine P en amont (aspartate 187) et en aval (isoleucine 215) du linker GSG. S'hybride sur la séquence en amont du résidu 187, le linker, et la séquence en aval du résidu 272 | 91 |
| TcCRTΔPterΔ288Reverse | AAA ATC CGG GTT CGG GAT ACC GCT GCC ATC CAC AAT CTC ACG CGG | | |
| EhCRTΔPbisΔ288Forward | GAC ATG CTG GCA CCG AAA GGC AGC GGT TAC GTG TAC GAC CCG GAG | Suppression de deux régions du domaine P en amont (lysine 188) et en aval (tyrosine 282) du linker GSG. S'hybride sur la séquence en amont du résidu 188, le linker, et la séquence en aval du résidu 282 | 87 |
| EhCRTΔPbisΔ288Reverse | CTC CGG GTC GTA CAC GTA ACC GCT GCC TTT CGG TGC CAG CAT GTC | | |
| EhCRTΔPterΔ288Forward | CCG AAA GAA ATT GAC GAT GGC AGC GGT ATT GCG AAC CCT GAC TAC | Suppression de deux régions du domaine P en amont (aspartate 192) et en aval (isoleucine 277) du linker GSG. S'hybride sur la séquence en amont du résidu 192, le linker, et la séquence en aval du résidu 277 | 92 |
| EhCRTΔPterΔ288Reverse | GTA GTC AGG GTT CGC ATT ACC GCT GCC ATC GTC AAT TTC TTT CGG | | |

TABLEAU 2.II.9 : OLIGONUCLEOTIDES UTILISES POUR INTRODUIRE DES DELETIONS DANS LE DOMAINE P DES CALRETICULINES DE *T. CRUZI* ET *E. HISTOLYTICA* (LA SEQUENCE EST DONNEE DE 5' EN 3')

II.4. Extraction de l'ADN plasmidique

La mini et la midi préparation sont utilisées pour extraire et purifier en grande quantité les ADN plasmidiques issus de 3 mL (mini) ou 300 mL (midi) de culture bactériennes. Les mini et midi préparations ont été effectuées selon les protocoles fournis par Nucleospin Plasmid® et NucleoBond® Xtra Midi de chez Macherey Nagel™, respectivement. Ces kits utilisent une extraction par lyse alcaline, suivie d'une précipitation des ADN chromosomiques, et des protéines. Une dernière étape de purification par chromatographie sur une colonne de silice permet d'obtenir l'ADN plasmidique pur. Après une concentration par précipitation, le culot contenant l'ADN plasmidique est lavé à l'éthanol 70%, séché et dissous dans de l'eau.

II.5. Quantification et contrôle de l'ADN

Les différents échantillons d'ADN (produits de PCR ou plasmides linéarisés) sont analysés sur gel d'agarose 1%. Pour cela les échantillons sont dilués avec du colorant de charge 5x (EDTA 0,1 M, bleu de bromophénol 0,25%, xylène cyanol 0,25%, glycérol 30%) puis déposés sur gel d'agarose 1% (p/v) contenant d'un agent intercalant (SYBR® Green, Invitrogen) (1/1000 V/V) permettant la révélation de l'ADN.

La migration s'effectue pendant 1 h à 100 V dans du tampon TAE (Tris-Acetate 40 mM EDTA 1 mM pH 8) et les fragments d'ADN sont révélés grâce à une lampe UV. Les différents fragments d'ADN linéaires migrent à travers le gel en fonction de leurs tailles et sont identifiés par comparaison avec des fragments d'ADN standard de référence, déposés dans les mêmes conditions (marqueurs de taille).

La concentration d'ADN double brin est obtenue en mesurant l'absorbance à 260 nm à l'aide d'un spectrophotomètre (nanodrop 2000 Thermo Scientific) selon la relation suivante : **1 unité d'absorbance à 260 nm correspond à 50 µg/mL** pour de l'ADN double brin. La pureté de l'ADN purifié est estimée en fonction de la valeur du rapport absorbance 260 nm/absorbance 280 nm, qui doit être supérieur à 2.

II.6. Séquençage

Pour vérifier l'intégrité des gènes codant pour les protéines d'intérêt dans les vecteurs et qu'aucune mutation, délétion ou insertion supplémentaire n'a été introduite lors des différents cycles de PCR, les séquences des gènes codant pour les protéines d'intérêt sont systématiquement vérifiées par séquençage. Pour cela les ADN plasmidiques purifiés sont envoyés sous forme soluble à la société GATC Biotech aux concentrations recommandées. La société séquence l'échantillon sur le principe de la méthode de Sanger : l'ADN est répliqué à partir d'une amorce complémentaire à une séquence placée en amont (sens) ou en aval (antisens) du gène d'intérêt. Dans notre cas une amorce complémentaire au promoteur T7 est utilisée pour le séquençage « sens » des différentes constructions plasmidiques. Une amorce BGHREV complémentaire du site de polyadénylation de l'hormone de croissance bovine est utilisée pour le séquençage « anti-sens » du plasmide pcDNA3.1.

II.7. Transformation des cellules bactériennes

Les cellules *E. coli* BL21 (DE3) sont rendues compétentes en utilisant le protocole suivant: la culture bactérienne est stoppée en phase exponentielle puis traitée par une solution CH₃COOK 30 mM, CaCl₂ 10 mM, KCl 100 mM, MnCl₂ 50 mM, Glycérol 15% pH 7 après centrifugation. Après incubation de 10 minutes dans la glace puis centrifugation, le culot est repris dans 4 mL de solution MOPS 10 mM, CaCl₂ 75 mM, KCl 10 mM pH 7. Ce traitement permet d'augmenter la perméabilité membranaire des cellules. Les bactéries sont transformées par la méthode établie (Sambrook et al., 1989) : 1 ng de plasmide est mélangé avec 50 µL de bactéries compétentes. Après une incubation de 30 minutes dans la glace, le mélange subit un choc thermique de 45 s à 42°C. Puis, après une seconde incubation dans la glace de 2 minutes, les cellules sont incubées dans 500 µL de milieu LB pendant 1 h à 37°C sous agitation (220 rpm). Ensuite les cellules sont étalées sur une boîte de Pétri contenant du milieu LB solide et de la kanamycine 1 mM et incubées à 37°C sur la nuit.

Les cellules *E. coli* XL10-Gold ultracompetentes utilisées pour générer les différents fragments de calréticuline par PCR et fournies avec le kit QuickChange II XL Site-Directed Mutagenesis (Agilent technologies) ont été transformées selon le protocole du fournisseur. Un volume de 45 µL de cellules est incubé 10 minutes dans un tube de 14 mL (BD Falcon polypropylène) dans la glace en présence de 2 µL de β-mercaptoéthanol. Puis un volume de 2 µL d'ADN prétraité avec DpnI est mélangé aux cellules. Après une incubation de 30 minutes dans la glace, le mélange subit un choc thermique de 30 s à 42°C. Puis, après une seconde incubation dans la glace de 2 minutes, les cellules sont incubées dans 500 µL de milieu NZY⁺ préchauffé à 42°C pendant 1 h à 37°C sous agitation (250 rpm). Ensuite les cellules sont étalées sur une boîte de Pétri contenant du milieu LB solide et de la kanamycine 1 mM et incubées à 37°C sur la nuit.

III. Production des protéines recombinantes

La liste des protéines utilisées au cours de cette thèse et leurs paramètres physicochimiques sont présentés dans l'annexe 4, page D.

III.1 Principe de la surproduction de protéines recombinantes

La surproduction d'une protéine recombinante repose sur notre capacité à contrôler et induire la synthèse en grande quantité de cette protéine, en maîtrisant notamment la transcription de l'ADN en ARNm. Pour cela le gène codant pour la protéine d'intérêt est introduit dans un plasmide d'expression en aval d'un promoteur transcriptionnel inductible ou possédant une importante activité constitutive. Les plasmides sont ensuite insérés dans différents types cellulaires possédant toute la machinerie transcriptionnelle et traductionnelle.

Au cours de ma thèse deux systèmes d'expression ont été utilisés pour permettre une surproduction des protéines recombinantes d'intérêt en cellules de mammifère et en cellules bactériennes.

Après transformation des cellules *E. coli* BL21 (DE3) avec les plasmides pJexpress411:73151 (DNA2.0) (**Figure 2.I.1**) contenant les gènes codant pour les différents fragments de calréticuline, l'induction repose sur les mécanismes mis en oeuvre par la bactérie pour réguler l'utilisation des sources d'énergie carbonées du milieu de culture. La première phase de la croissance permet d'obtenir une biomasse conséquente, puis l'ajout d'IPTG, un analogue du lactose, permet de lever la répression du promoteur *lac* dont dépend l'expression de la T7 polymérase et donc des gènes clonés sous le contrôle du promoteur T7 (**Figure 2.III.1 et 2.III.2**).

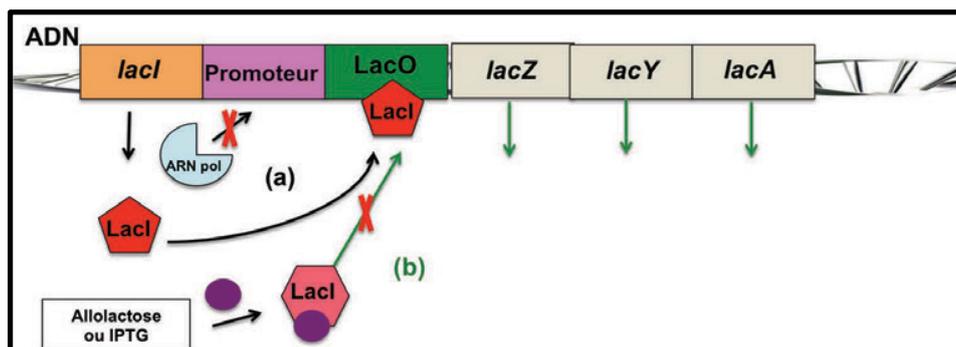


FIGURE 2.III.1 : REGULATION DE L'OPÉRON LACTOSE

L'opéron lactose est composé d'un promoteur (P), d'un opérateur (O) et de trois gènes *lac Z*, *lac Y* et *lac A* codant respectivement pour la β -galactosidase, la lactose perméase et la thiogalactoside transacétylase. En amont de cet opéron se trouve le gène *lac I* codant pour une protéine Lac I. En absence de lactose dans la cellule, la protéine Lac I s'organise en homo tétramère et se fixe sur l'opérateur réprimant la transcription en empêchant la fixation de l'ARN polymérase sur le promoteur (a). Lors de l'entrée du lactose dans la cellule, celui-ci est transformé en allolactose. Ce dernier, comme son analogue l'IPTG, se fixe alors sur le répresseur (tétramère Lac I) et modifie sa conformation, l'empêchant ainsi de se lier à l'opérateur et donc de réprimer la transcription (b).

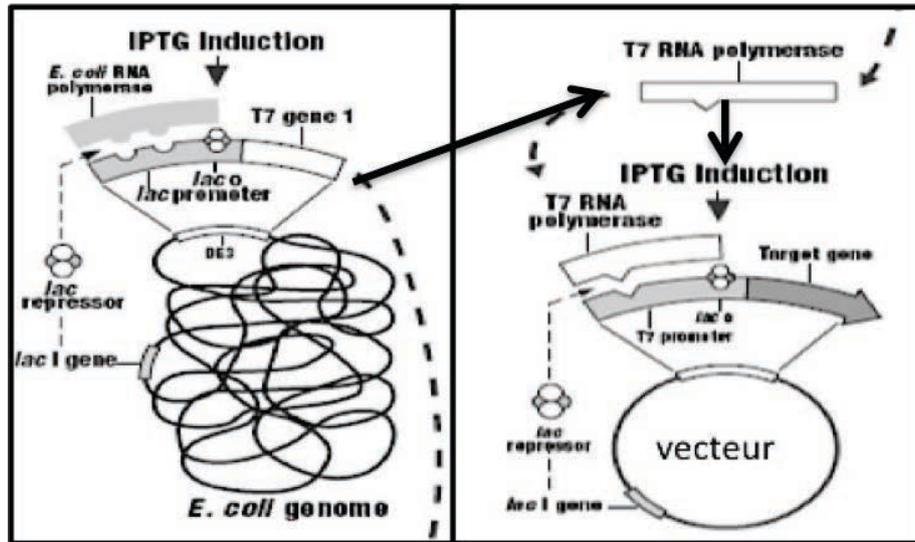


FIGURE 2.III.2: PRINCIPE DE L'EXPRESSION D'UNE SEQUENCE CODANTE CLONÉE DANS UN VECTEUR UTILISANT UN PROMOTEUR T7 SOUS LE CONTRÔLE DE L'OPÉRON LACTOSE (ISSUE DU MANUEL NOVAGEN)

La souche utilisée C41 (DE3) possède le gène codant pour l'ARN polymérase T7, sous le contrôle de l'opérateur lacO de l'opéron lactose. Ainsi, l'utilisation de l'IPTG (analogue du lactose) permet de lever la répression sur l'opérateur et la synthèse de l'ARN polymérase T7, celle-ci pouvant ainsi se fixer au promoteur T7 en amont du gène d'intérêt sur le vecteur, ce promoteur T7 également fusionné à l'opérateur lacO dont la répression a été levée par l'entrée de l'IPTG.

Dans le cas d'une expression en cellules de mammifères, les cellules HEK293 ont été transfectées avec le plasmide pcDNA3.1 contenant le gène codant pour les régions globulaires de C1q sous contrôle du promoteur CMV (promoteur cytomégalovirus reconnu par l'ARN polymérase III des cellules de mammifères) (**Figure 2**). Contrairement au promoteur T7, le promoteur CMV n'est pas inductible, mais possède une forte activité transcriptionnelle constitutive. Grâce à une pression de sélection apportée par l'ajout d'un antibiotique spécifique, le plasmide est intégré au génome des cellules et sera répliqué et transcrit en même temps que l'ADN génomique des cellules.

III.2. Protocole de surproduction des fragments de calréticuline

Afin de surproduire les différents fragments de calréticuline de *T. cruzi* et *E. histolytica*, un protocole unique et préalablement optimisé a été utilisé. Pour préparer 1 L de culture, 50 mL de milieu LB contenant de la kanamycine 1 mM sontensemencés à partir d'un stock glycérol de cellules *E. coli* BL21 (DE3) transformées avec les plasmides pJexpress411:73151-CRTs, et incubées à 37°C sous agitation (220 rpm) sur la nuit. Le lendemain 1 L de milieu LB

contenant de la kanamycine 1 mM estensemencé avec 10 mL de la pré-culture et incubé à 30°C sous agitation (220 rpm). Lorsque l'absorbance mesurée à 600 nm atteint une valeur comprise entre 0,6 et 0,8, la culture est incubée à 4°C pendant une heure. La production du fragment de calréticuline est alors induite par ajout d'IPTG à une concentration finale de 1 mM et la culture est incubée à 20°C sous agitation pendant 20 h.

III.3. Protocole de surproduction des fragments de calréticuline pour la RMN

Pour réaliser une expérience de RMN, il est nécessaire que la protéine étudiée soit marquée avec des isotopes du carbone et de l'azote magnétiquement actifs ^{13}C et ^{15}N . Pour cela nous sommes dans l'obligation de travailler avec un milieu de culture dont la composition chimique est maîtrisée, ici le milieu M9 enrichi décrit dans le **tableau 2.I.2**. Dans ces conditions, une pré-culture de 5 mL de milieu LB contenant de la kanamycine 1 mM estensemencée à partir d'une colonie de cellules *E. coli* BL21 (DE3) transformées avec les plasmides pJexpress411:73151-CRTs et incubées à 37°C sous agitation (220 rpm) pendant 9 h.

Une seconde pré-culture de 100 mL de milieu M9 enrichi ^{15}N ou $^{13}\text{C}/^{15}\text{N}$ contenant de la kanamycine 1 mM est inoculée avec 1mL de la pré-culture LB, puis incubée à 37°C sous agitation (220 rpm) sur la nuit.

Le lendemain, après avoir vérifié que l'absorbance à 600 nm de la pré-culture est supérieure à 1, cette dernière est centrifugée à 4000 rpm pendant 20 minutes. Le culot cellulaire est alors repris dans 1 L de milieu M9 enrichis en ^{15}N ou $^{13}\text{C}/^{15}\text{N}$ (l'absorbance à 600 nm doit être aux alentours de 0,2) en présence de kanamycine 1 mM et la culture est incubée à 37°C sous agitation (220 rpm). Lorsque l'absorbance mesurée à 600 nm atteint une valeur comprise entre 0,6 et 0,8, la culture est incubée à 4°C pendant une heure. La production du fragment de calréticuline est alors induite par ajout d'IPTG à une concentration finale de 1 mM et la culture est incubée à 20°C sous agitation pendant 20 h.

III.4. Protocole de surproduction des GRC1q recombinantes

Les cellules HEK 293-FTM (Life Sciences Technologies) sont transfectées avec l'ADN plasmidique pcDNA3.1-GRC1q en présence de 293FectinTM (Invitrogen) selon les recommandations du fournisseur. Un volume de 30 mL de cellules à 1.10^6 cellules/mL est

transfecté avec 30 µg de plasmide en présence de 60 µL d'agent de transfection. Les cellules sont cultivées à une température de 37°C sous agitation (140 rpm) sous une atmosphère chargée de 5% de CO₂ dans le milieu sans sérum FreeStyle 293 (Life Sciences Technologies) en absence d'antibiotique.

Après 72 h, pour établir un lignée stable de cellules ayant intégré le plasmide pcDNA3.1-GRC1q à leur génome pour produire en permanence la protéine d'intérêt, les cellules HEK 293-FTM sont maintenues en suspension à une concentration de 0,8.10⁶ cellules/mL dans le milieu sans sérum FreeStyle 293. Les cellules sont cultivées à une température de 37°C sous agitation constante (140 rpm) sous une atmosphère chargée de 5% de CO₂ et en présence de néomycine (Geneticin[®]) à une concentration de 400 µg/mL.

Le milieu de culture et l'antibiotique sont renouvelés toutes les 72 h afin de maintenir une concentration de cellules vivantes comprise entre 0,6 et 0,8.10⁶ cellules/mL dans un volume de 250 mL. Lorsque le nombre de cellules mortes dues à la pression de sélection descend en dessous de 10% des cellules totales, c'est à dire qu'au moins 90% des cellules ont intégré le plasmide pcDNA3.1-GRC1q à leur génome, ces dernière produisent alors en permanence la protéine recombinante d'intérêt qui est exportée dans le surnageant. Il est alors possible de congeler les cellules pour une future production.

Les cellules HEK 293-FTM sont maintenues en suspension à une concentration comprise entre 0,6 et 0,8.10⁶ cellules/mL dans 250 mL de milieu sans sérum FreeStyle 293 renouvelé toutes les 72 h et cultivées à une température de 37°C sous agitation constante (140 rpm) sous une atmosphère chargée de 5% de CO₂ et en présence de néomycine (Geneticin[®]) à une concentration de 400 µg/mL. A chaque renouvellement du milieu les 250 mL de surnageant contenant la protéine GRC1q recombinante sont conservés.

IV. Purification de protéines

IV.1. Principe de la purification de protéines

Les protéines recombinantes exprimées en cellules bactériennes sont généralement exportées dans le cytoplasme ou le périplasme de ces cellules. La rupture de la paroi et de la membrane cellulaire des bactéries entraîne la libération de l'intégralité du contenu cytoplasmique : les organites, acides nucléiques, macromolécules solubles, protéines endogènes, protéine recombinante d'intérêt, et nous sommes en présence des débris cellulaires (fragments de

membrane et protéines associées). De la même manière le milieu de culture des cellules de mammifères contient les différentes protéines sécrétées, et les constituants chimiques du milieu.

Pour les caractérisations structurales et fonctionnelles des protéines étudiées au cours de cette thèse, il est nécessaire de posséder un échantillon pur dans un tampon dont la composition chimique est maîtrisée. En effet, la présence de protéines contaminantes, ou la non maîtrise de la nature des sels et du pH de la solution peut engendrer de nombreux problèmes. D'un point de vue fonctionnel, les contaminants protéiques peuvent fausser le calcul de la concentration, inhiber, modifier ou gêner l'activité de la protéine étudiée (modification de la conformation, encombrement stérique au niveau d'un site actif ou d'interactions). D'un point de vue structural, être en possession d'un échantillon pur et homogène est également essentiel, les contaminants en trop grande quantité ou des formes oligomériques diverses de la protéine d'intérêt peuvent perturber la formation des cristaux ou aboutir à des cristaux de mauvaise qualité (mosaïcité, faible pouvoir de diffraction). La majorité des techniques utilisées pour les études structurales (que je décrirai plus en détails dans la section V.) donnent une image moyenne de toutes les molécules présentes dans l'échantillon.

Pour purifier les protéines, la méthode la plus couramment utilisée est la chromatographie. La technique repose sur la force d'adsorption d'une phase mobile liquide (solution protéique) sur une phase stationnaire solide (résine constituée de billes de silice ou sépharose fonctionnalisées avec des groupements chimiques ou des ligands). Plus l'adsorption des molécules sur la phase fixe est forte, plus elles seront éluées lentement à travers la colonne.

Cette technique permet de séparer les molécules par élution différentielle des solutés en fonctions de leurs propriétés physico-chimiques et fonctionnelles :

- Taille et état d'oligomérisation (chromatographie d'exclusion de taille).
- Charges électrostatiques de surface (chromatographie d'échange d'anions ou de cations).
- Hydrophobicité (chromatographie d'interactions hydrophobes, chromatographie en phase inverse).
- Affinité pour un ligand ou un ion métallique (chromatographie d'affinité, chromatographie d'affinité d'ions métalliques (IMAC)). (**Figure 2.IV.1**)

Il existe d'autres techniques ne reposant pas sur l'utilisation d'une phase solide mais reposant tout de même sur les propriétés physico-chimiques des protéines, comme la solubilité relative

dans un solvant (précipitation fractionnée au sulfate d'ammonium) ou résistance à des conditions extrêmes de températures (chauffage au bain-marie).

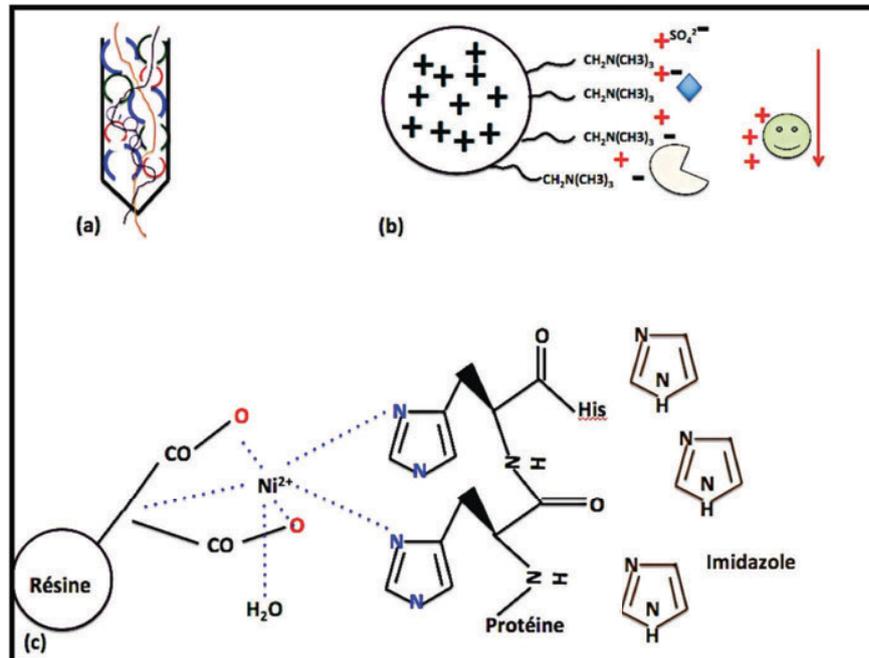


FIGURE 2.IV.1: PRINCIPE DES DIFFERENTES CHROMATOGRAPHIES EN PHASE LIQUIDE UTILISEES POUR LA PURIFICATION DES PROTEINES RECOMBINANTES D'INTERET

(a) Les colonnes de chromatographie d'exclusion de taille sont composées d'une résine constituée de billes microporeuses de tailles différentes. Les protéines les plus petites traversent la totalité de ces billes et sont ainsi retardées, tandis que les molécules les plus importantes sont exclues de ces billes et seront donc éluées les premières. Cette technique permet de séparer les protéines en fonction de leur taille et donc de séparer les différentes formes oligomériques ou les agrégats d'une même protéine.

(b) La chromatographie échangeuse d'ions repose sur l'utilisation d'une résine fonctionnalisée avec des groupements chimiques chargés. Ici est représenté le cas d'une échangeuse d'anions, la résine chargée positivement retient les ions négatifs, ainsi que les protéines possédant des groupements chargés négativement à leur surface. Plus ces groupements sont nombreux et accessibles, plus l'adsorption est forte. Le passage d'un tampon de faible force ionique à un tampon avec une importante concentration en sels entraîne une compétition entre les ions apportés en excès par le tampon et les charges de surface de la protéine, permettant son élution. Le même principe s'applique aux chromatographies d'interactions hydrophobes, utilisant une phase solide décorée de groupements hydrophobes.

(c) Les résines utilisées pour les chromatographies d'affinité sont fonctionnalisées avec un ligand interagissant fortement et de manière spécifique avec la protéine d'intérêt. Cette interaction peut être naturelle (anticorps, ligand physiologique) ou il est possible de la créer artificiellement en ajoutant une protéine ou un groupement chimique spécifique à la protéine d'intérêt lors du clonage. Dans ce projet, une étiquette composée de 6 histidines a été ajoutée lors du clonage. Les noyaux imidazoles des histidines chélatent l'ion nickel, permettant de fixer la protéine. L'ajout d'imidazole en solution entraîne une compétition avec l'étiquette histidine et une élution de la protéine.

IV.2. Protocole de purification des fragments de calréticuline

Les différents fragments des calréticulines de *T. cruzi* et *E. histolytica* ont été purifiés en utilisant les protocoles établis lors des travaux menés au laboratoire sur la calréticuline humaine (Chouquet et al., 2011).

IV.2.a Extraction par lyse des cellules

Après centrifugation des cellules à 5000 rpm pendant 20 min, les culots sont repris dans du **tampon de lyse (Tris 50 mM, NaCl 300 mM, CaCl₂ 5 mM, Imidazole 10 mM, Glycérol 10%, pH 7,5)**, à raison de 30 mL pour un culot cellulaire correspondant à 1 L de culture, et les cellules sont lysées à froid en présence de DNase 2 µg/mL et d'inhibiteurs de protéase (cOmplete, EDTA-free ROCHE) par des ultrasons générés à 20 kHz par un sonicateur (Vibra cell, Sonics & Materials, INC.) à une puissance de 75 W pendant 12 minutes par période de 1 s d'ultrasons puis 5 s d'arrêt (soit 2 minutes de lyse). Après centrifugation du lysat à 15000 rpm pendant 45 min, le surnageant de sonication est prélevé.

IV.2.b Chromatographie d'affinité d'ion métallique (IMAC)

Le surnageant de sonication est injecté sur une colonne d'affinité d'ion métallique de type Histrap FastFlow 5 mL (GE Healthcare Life Sciences), matrice d'agarose sur laquelle sont greffés des groupements acide iminodiacétiques chélateurs du nickel, couplée à un système FPLC (Äkta Purifier, GE Healthcare Life Sciences). La colonne est préalablement équilibrée avec du **tampon de lyse (Tris 50 mM, NaCl 300 mM, CaCl₂ 5 mM, Imidazole 10 mM, Glycérol 10%, pH 7,5)**. Dans ces conditions, les différents fragments de calréticuline sont adsorbés à la surface de la colonne grâce à la présence de l'étiquette histidine. Après un lavage de la colonne avec 5 volumes de colonne de tampon, l'élution s'effectue par un gradient linéaire croissant d'imidazole de **10 mM à 500 mM** dans le même tampon **sur 50 min sous un débit de 1 mL/min**. Les fragments de calréticuline sont élués à une concentration d'imidazole de 150 mM.

IV.2.c Chromatographie d'exclusion de taille

Les fractions issues de la chromatographie IMAC contenant la protéine sont concentrées à un volume correspondant à 5% du volume total de la colonne de chromatographie d'exclusion de taille et en respectant les limites de solubilité des protéines. Les échantillons sont alors injectés sur une colonne de chromatographie d'exclusion de taille HiLoad 16/60 Superdex 75

prep grad (120 mL) ou HiLoad 26/60 Superdex 75 prep grad (320 mL) dont la résine est constituée d'agarose et de dextran sous forme de billes (22 à 44 µm de diamètre) microporeuses, couplée à un système FPLC (Äkta Purifier, GE Healthcare Life Sciences). Les colonnes sont préalablement équilibrées avec le **tampon de GF Tris 20 mM, NaCl 150 mM, CaCl₂ 5 mM, pH 8**. Le domaine de fractionnement de cette résine se situe entre 3 et 70 kDa. L'élution s'effectue selon le principe présenté **Figure 2.IV.1**, à un débit de 1,5 mL/min (16/60) ou 2,5 mL/min (26/60) sur un volume total de 1,5 volume colonne (180 et 480 mL respectivement).

Les échantillons peuvent également être injectés sur une colonne de chromatographie d'exclusion de taille TSKgel G3000SW (Tosoh Bioscience LLC), (21,5 mm ID x 30 cm, 109 mL) dont la résine est constituée de billes de silice (13 µm) microporeuse (0,25 µm) couplée à une colonne de garde composée de la même résine (21,5 mm ID x 1,5 cm). Le domaine de fractionnement de cette résine se situe entre 10 et 500 kDa. L'élution s'effectue à un débit de 5 mL/min sur un volume total de 150 mL.

IV.2.d Concentration des fragments de calréticuline

Les fractions contenant les fragments de calréticuline sont identifiées par gel SDS-PAGE (**voir IV.4**) et rassemblées. Les échantillons sont concentrés par centrifugation en utilisant un centricon AMICON Ultra-15 (Millipore), avec une membrane en cellulose régénérée (seuil de coupure 10 kDa) à une concentration de comprise entre 3 et 20 mg/mL selon les utilisations et conservés à 4°C ou -80°C.

La HuCRT utilisée pendant cette thèse a été purifiée au laboratoire par Anne Chouquet en suivant le même protocole.

IV.3. Protocole de purification de C1q et des GRC1q dérivées du sérum et recombinantes

La protéine C1q, ainsi que les régions globulaires GRC1q ont été purifiées à partir du sérum humain en utilisant les protocoles établis par Gérard Arlaud, Nicole Thielens et leurs collaborateurs (**Arlaud et al., 1979**).

IV.3.a. Extraction plasmatique

IV.3.a.1. Capture du complexe C1 sur des complexes immuns

Deux litres de plasma recalcifié (sérum) sont centrifugés à 9500 rpm pendant 20 min. Le surnageant est mélangé avec un inhibiteur de protéase (SBTI 50 µg/mL) et 2 g de complexes immuns (préparés en mélangeant à équivalence de l'ovalbumine et des Ig antiovalbumine) et incubé à 0°C pendant 45 min. Après centrifugation à 7500 rpm pendant 7,5 min, le culot est re suspendu dans 400 mL de **tampon de lavage des complexes immuns (Tris 20 mM, NaCl 120 mM, CaCl₂ 5 mM, pH 7) + SBTI 40 µg/mL** (le cycle centrifugation-lavage est reproduit 3 fois).

Ensuite le culot est repris dans 300 mL de **tampon de lavage des complexes immuns** sans SBTI et incubé 35 min à 30°C pour activer les protéases. A ce stade le complexe C1 (composé des protéines C1q, C1r et C1s) est fixé par sur les complexes immuns par l'intermédiaire de C1q.

IV.3.a.2. Dissociation du complexe C1

Pour obtenir la protéine C1q pure, il est essentiel de dissocier le complexe C1. Après l'incubation de 35 min à 30°C, la solution est centrifugée à 7500 rpm pendant 7,5 min et le culot est repris dans 100 mL de **tampon d'extraction r+s (Tris 50 mM, NaCl 60 mM, EDTA 20 mM, pH 7)** (l'association des protéases C1r-C1s à C1q étant dépendante du calcium, la capture de ce dernier par l'EDTA entraîne la dissociation) et centrifugé à 7500 rpm pendant 7,5 min (opération répétée 2 fois). Les deux surnageants (contenant C1r et C1s) sont conservés pour une future purification des protéases. A cette étape, C1q est toujours fixé aux complexes immuns dans le culot.

IV.3.a.3. Purification de C1q

Le culot est repris dans 100 mL de **tampon d'extraction de C1q (Tris 50 mM, NaCl 700 mM, pH 10)** (permettant la dissociation de C1q des complexes immuns) et centrifugé à 8000 rpm pendant 8 min (2 fois). Les 200 mL de surnageant sont dialysés sur la nuit contre 5 L de **tampon de dialyse de C1q (acétate de Sodium 230 mM, pH 5,2)**.

Après dialyse, le surnageant est déposé à un débit de 45 mL/h sur une colonne échangeuse de cations de type CM52 (70 mL) (Whatman), matrice de cellulose sur laquelle sont greffés des groupements carboxyméthyle éther chargés négativement. La colonne est équilibrée avec 1 L

de **tampon de dialyse de C1q**. A ce pH, C1q, dont le pI est 8,84, est chargé positivement et s'adsorbe à la surface de la colonne. Après lavage de la colonne jusqu'à atteindre une absorbance à 280 nm inférieure à 0,02, l'éluion se fait par un gradient linéaire croissant de NaCl de 0 à 200 mM dans le même tampon, sur 1 L sous un débit de 50 mL/h, avec collecte par fraction de 4,5 mL.

IV.3.a.4. Concentration de C1q

Les fractions contenant C1q sont identifiées par gel SDS-PAGE et Western Blot (**voir IV.4**) et rassemblées. Une partie est dialysée contre 2 L de **tampon de conservation de C1q (acétate de Sodium 230 mM, NaCl 300 mM, pH 5,2)**, puis concentrée en utilisant une cellule de concentration AMICON, avec une membrane en cellulose régénérée (seuil de coupure 30 kDa) à une concentration de 1 mg/mL et conservée à -20°C.

La seconde partie de l'échantillon est utilisée pour la purification des régions globulaires de C1q (GRC1q).

IV.3.a.5. Digestion des régions collagènes de C1q

Pour isoler les régions globulaires de C1q, les régions collagènes doivent être éliminées par digestion enzymatique.

La fraction de C1q est dialysée contre **le tampon des GRC1q (Tris 50 mM, NaCl 250 mM, CaCl₂ 5 mM, pH 7,4)** et mélangée avec de la collagénase (collagénase de *Clostridium histolyticum*, Sigma-Aldrich), (3000 U/mg de C1q). La digestion s'effectue à 37°C pendant 24 h.

IV.3.a.6. Séparation des fragments collagènes et des régions globulaires.

Après la digestion, l'échantillon est filtré sur un filtre millipore 0,45 µm et injecté (par fractions de 2 mL maximum) sur une colonne de chromatographie d'exclusion de taille TSKgel G3000SW (Tosoh Bioscience LLC) (**voir section IV.2.3**) équilibrée avec **le tampon des GRC1q (Tris 50 mM, NaCl 250 mM, CaCl₂ 5 mM, pH 7,4)**. L'éluion s'effectue selon le principe présenté **Figure 2.IV.1**, à un débit de 5,8 mL/min sur un volume total de 200 mL, avec collecte de fractions 6mL.

IV.3.a.7. Concentration des GRC1q

Les fractions contenant les GRC1q sont identifiées par gel SDS-PAGE et Western Blot (**voir IV.4**) et rassemblées. L'échantillon est concentré en utilisant une cellule de concentration AMICON, avec une membrane en cellulose régénérée (seuil de coupure 10 kDa) à une concentration de 1 mg/mL et conservée à -20°C.

De la même manière, les régions collagènes de C1q (CLR) sont purifiées par une digestion contrôlée à la pepsine, puis séparées des régions globulaires par chromatographie d'exclusion de taille (**Tacnet-Delorme et al., 2001**). Pour les expériences SPR, le récepteur gC1qR nous a été gracieusement fourni par le docteur Berhane Ghebrehwet de l'université Stony Brook (Stony Brook New-York USA 11794-8161)

IV.3.b. GRC1q exprimée s en cellules de mammifères

Le surnageant de culture des cellules HEK293-F contenant les GRC1q recombinantes (500 mL) est dialysé contre 15 L (3 bains de 5 L) de **tampon léger (MES 50 mM, NaCl 25 mM, pH 6,5)** en présence d'inhibiteur de protéase DFP 1 mM final.

Après dialyse le surnageant est injecté à un débit de 90 mL/h sur une colonne échangeuse de cation de type SP-Sépharose Fast Flow (50 mL) (GE Healthcare Life Sciences), matrice d'agarose sur laquelle sont greffés des groupements sulfopropyles chargés négativement. La colonne est équilibrée avec 1 L de **tampon léger**. A ce pH, les GRC1q, dont le pI est 8,36 sont chargés positivement et s'adsorbent à la surface de la colonne.

Après lavage de la colonne avec 2 L de **tampon léger**, l'élution se fait par un gradient linéaire croissant de NaCl de 25 à 500 mM dans le même tampon sur 1 L sous un débit de 90 mL/h. La collecte est faite par fractions de 5 mL.

Les fractions contenant les GRC1q recombinantes sont identifiées par gel SDS-PAGE et Western Blot (**voir IV.4**) et rassemblées. L'échantillon est concentré par centrifugation en utilisant un centricon AMICON Ultra-15 (Millipore), avec une membrane en cellulose régénérée (seuil de coupure 10 kDa) à une concentration de 1 mg/mL et conservé à 4°C.

IV.4. Analyse des échantillons protéiques

IV.4.a. Quantification de la concentration en protéine

La concentration des protéines purifiées est calculée en mesurant l'absorbance à une longueur de 280 nm avec un spectrophotomètre à nanogouttes (nanodrop 2000, Thermo Scientific) ou à cuves (Biophometer, Eppendorf). Les protéines absorbent fortement à cette longueur d'onde grâce à la présence des résidus aromatiques phénylalanine, tryptophane et tyrosine

La concentration molaire est déterminée à partir de l'absorbance mesurée en utilisant la loi de Beer-Lambert (*Beer, 1852*) :

$$A = \epsilon \cdot l \cdot C$$

Avec A l'absorbance à 280 nm de l'échantillon, ϵ le coefficient d'extinction molaire à 280 nm en $M^{-1} \cdot cm^{-1}$, l la longueur du trajet optique en cm et C la concentration de l'échantillon en $mol \cdot L^{-1}$

De la même manière les concentrations massiques sont calculées à partir de la concentration molaire en multipliant cette dernière par la masse moléculaire de l'échantillon (M).

$$\text{Ainsi } C \text{ (g. L}^{-1}\text{)} = C \text{ (mol.L}^{-1}\text{)} \times M_{\text{échantillon}} \text{ (g.mol}^{-1}\text{)}$$

Les coefficients d'extinction molaire de chaque échantillon sont calculés par le programme ProtParam disponible sur le serveur Expasy (<http://web.expasy.org/protparam/>). Ce programme se base sur la méthode d'Edelhoch (**Edelhoch, 1967**) en utilisant les coefficients d'extinction molaire de la tyrosine et du tryptophane (**Pace et al., 1995**), ainsi que la participation des cystéines sous forme de pont disulfure.

$$\epsilon_{\text{(protéine)}} = \text{Nombre Tyr. } \epsilon_{\text{(Tyr)}} + \text{Nombre Trp. } \epsilon_{\text{(Trp)}} + \text{Nombre Cys. } \epsilon_{\text{(cys)}}$$

$$\text{Avec } \epsilon_{\text{(Tyr)}} = 1490 M^{-1} \cdot cm^{-1}, \epsilon_{\text{(Trp)}} = 5500 M^{-1} \cdot cm^{-1} \text{ et } \epsilon_{\text{(Cys)}} = 125 M^{-1} \cdot cm^{-1}$$

IV.4.b. Electrophorèse sur gel de polyacrylamide (SDS-PAGE)

Pour identifier les fractions contenant les protéines d'intérêt collectées lors des différentes purifications, ainsi que pour contrôler la qualité de ces protéines (présence de contaminants, de fragments de dégradation, d'oligomères forts) et estimer la quantité, les différents échantillons sont analysés sur gel SDS PAGE (SDS PolyAcrylamide Gel Electrophoresis).

Cette méthode consiste à faire migrer un échantillon protéique chargé négativement à travers un maillage plus ou moins étroit selon le degré de polymérisation du gel en appliquant un courant électrique. Les protéines progressent donc à travers le gel dans un temps donné en fonction de leurs masses apparentes (influence du repliement bidimensionnel).

Le gel de polyacrylamide est composé de deux gels distincts (pH et maillages différents).

- **Un gel de concentration (stacking gel)** composé de : acrylamide–bis acrylamide (5%) (ratio 37 :1 Euromedex), Tris 125 mM pH 6,8, SDS 0,1%, persulfate d'ammonium 0,1% et N,N,N,N-tétra-méthyl-éthylènediamine (TEMED) 0,1%.
- **Un gel de séparation (running gel)** composé de : acrylamide–bis acrylamide (6 à 15%) (ratio 37 :1 Euromedex), Tris 390 mM pH 8,8, SDS 0,1%, persulfate d'ammonium 0,2% et N,N,N,N-tétra-méthyl-éthylènediamine (TEMED) 0,2%.

Les protéines analysées sont tout d'abord dénaturées et chargées négativement par un traitement thermique et chimique, leur donnant la capacité de toutes migrer dans la même direction (vers l'anode) et uniquement en fonction de leur masse apparente à travers le gel. Lorsque le volume de l'échantillon est trop important pour être déposé sur gel, celui-ci est précipité par un volume équivalent d'acide trichloroacétique 50% (TCA) pendant une nuit à 4°C. L'échantillon est ensuite centrifugé pendant 30 minutes à 13200 rpm, et le culot est lavé deux fois avec une solution d'acétone pure. Le culot est ensuite resuspendu dans un volume adapté de tampon de dénaturation.

Les protéines recombinantes issues de culture bactérienne sont dénaturées par chauffage à 95°C pendant 5 minutes dans une solution de **Tris 25 mM, SDS 1%, glycérol 4% bleu de bromophénol pH 6,8** en présence ou non de DTT 100 mM si nous souhaitons réduire les échantillons. Dans les deux cas de figure, les échantillons sont chauffés à 95°C pendant 5 min.

Certaines protéines, comme celles issues du plasma ne supportent pas une température de 95°C. Ces échantillons sont incubés à 37°C dans une solution de **Tris 100 mM, urée 4 M, SDS 1%, iodoacétamide 10 mM bleu de bromophénol pH 8 pendant 1 h**. Les échantillons peuvent également être réduits en étant incubés à 37°C pendant 1 h dans une solution de **Tris**

100 mM, urée 4 M, SDS 1%, DTT 50 mM pH 8 avant d'être alkylés par de iodoacétamide 140m M pendant 20 minuts à 37°C.

Ces traitements dénaturent les protéines grâce à la chaleur, l'urée et un détergent tensioactif, le SDS, ce dernier ayant également la particularité de se fixer à la molécule (1 molécule de SDS tous les deux acides aminés) apportant une charge négative à toutes les molécules.

Le gel est placé verticalement dans un tampon de migration (**Tris 25 mM, Glycine 192 mM, SDS 0,1%, pH 6,8-8,8**) grâce au système d'électrophorèse (Mighty Small II, Amersham Biosciences). Les échantillons sont déposés dans des puits au sommet du gel et un courant est appliqué par un générateur (Electrophoresis Power supply EPS 301, Amersham Biosciences) permettant aux protéines chargées négativement de migrer vers l'anode (**Figure 2.IV.3**).

Un mélange de protéines de taille connue (Prestained Protein Ladder, Euromedex) est déposé en parallèle des échantillons et utilisé comme standard pour estimer les masses des protéines d'intérêt.

A la fin de la migration les protéines présentes sur le gel sont révélées grâce à une coloration au bleu de Coomassie ou transférées sur une membrane de nitrocellulose afin de réaliser une expérience d'immunorévélation.

IV.4.c. Transfert sur membrane de nitrocellulose et immunorévélation (Western Blot)

Le transfère des protéines contenues dans le gel SDS PAGE sur une membrane de nitrocellulose permet de les identifier grâce à leur interaction avec des anticorps spécifiques.

Après la séparation des protéines sur gel SDS PAGE, ce dernier ainsi qu'une membrane de nitrocellulose, du papier filtre Whatman et des éponges sont incubés quelques minutes dans un **tampon de transfert (Tris 25 mM, glycine 192 mM, éthanol 20%)** et montés comme décrits dans la **figure 2.IV.2**. Ce montage est placé dans une cuve (Mini Blot Cell, Biorad) en présence du tampon de transfert. Le transfert des protéines s'effectue vers l'anode sur le même principe que le gel SDS PAGE en appliquant un courant électrique de 100 Volts à 4°C pendant 1 h ou de 12 V sur la nuit.

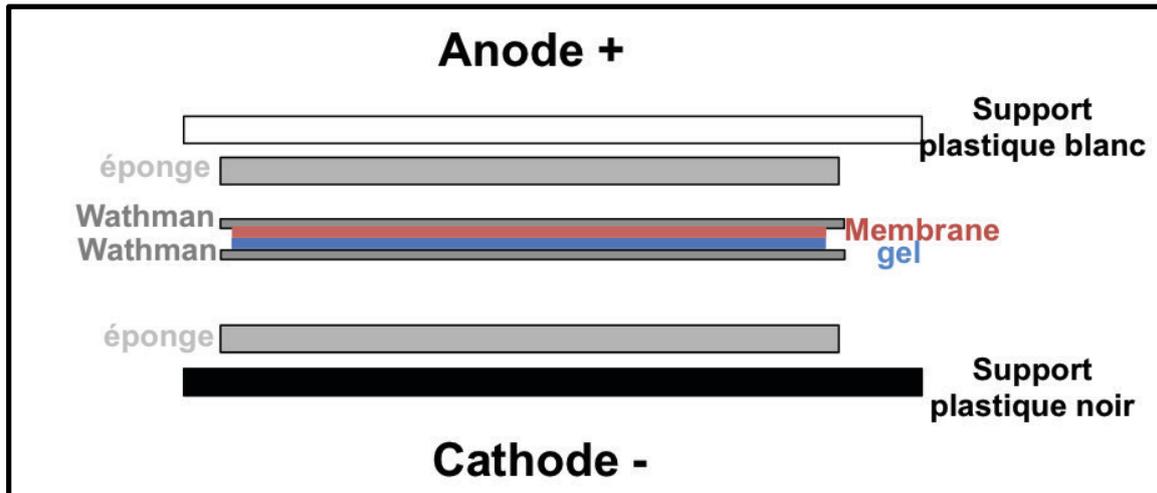


FIGURE 2.IV.2 : SCHEMA DE MONTAGE DE LA CASSETTE POUR LE TRANSFERT SUR MEMBRANE DE NITROCELLULOSE (WESTERN BLOT).

Après le transfert, la membrane est lavée par 3 bains successifs de 10 min de **tampon PBS** (NaCl 137 mM, KCl 2,7 mM, Na_2HPO_4 10 mM, KH_2PO_4 1,76 mM pH 7,4 Euromedex) et incubée dans une solution de blocage (**tampon PBS + 5% de lait lyophilisé**) pendant 1 h à 37°C sous agitation pour saturer les sites non spécifiques.

Un anticorps primaire dirigé spécifiquement contre un épitope de la protéine à identifier est dilué dans la solution de blocage à une concentration établie, puis la membrane est incubée en présence de l'anticorps sur la nuit à 4°C sous agitation.

Après 3 nouveaux lavages avec du tampon PBS + 0,05% de Tween 20 (PBS-T), la membrane est incubée dans du tampon de blocage avec un anticorps secondaire spécifique de l'isotype de l'anticorps primaire et couplé à la phosphatase alcaline pendant 1 h à 37°C.

Après les 3 derniers lavages avec du tampon PBS-T, la présence des protéines sur la membrane est révélée par ajout de 1 mL du substrat **NBT (nitro bleu tetrazolium chloride)/BCIP (5-Bromo-4-chloro-3-indolyl phosphate)** Uptima (Interchim) pendant quelques minutes. Le BCIP est déphosphorylé par la phosphatase alcaline et peut alors être oxydé par le NBT et former un précipité bleu. Quand le signal a l'intensité désirée la réaction est stoppée à l'eau.

IV.4.d. Spectrométrie de masse

La masse des protéines a été déterminée par spectrométrie de masse MALDI-TOF ou ESI sur la plateforme de spectrométrie de masse de l'ISGB par Luca Signor sur deux appareils

MALDI-TOF MS (Autoflex, Bruker Daltonics) et LC ESI-TOF MS (6210, Agilent Technologies).

Dans une expérience MALDI (ionisation par Désorption Laser Assistée par Matrice), les échantillons sont co-cristallisés avec une matrice d'acide sinapinique et ionisés (mono chargés) sous vide par un laser. Dans une expérience ESI (électronébulisation), les échantillons sont ionisés (multichargés) en solution à pression atmosphérique sous l'effet d'un champ électrique intense. Dans les deux cas les ions mono ou multichargés sont séparés et analysés dans un détecteur TOF (Temps de vol) qui permet de déterminer le rapport masse sur charge de l'ion (m/z) en fonction du temps mis par l'ion pour traverser l'analyseur. Un traitement informatique permet de déterminer la masse molaire expérimentale moyenne de la protéine avec une erreur de l'ordre de 0,1%.

IV.4.e. Séquençage N-terminal des protéines

La détermination de la séquence N-terminale des protéines a été réalisée par Jean-Pierre Andrieu sur la plateforme de séquençage de l'ISGB. La séquence est déterminée sur les 5 premiers résidus, longueur suffisante pour identifier précisément nos protéines et les éventuels contaminants. La technique se base sur la réaction de dégradation d'Edman (**Edman, 1949**). La protéine est mise en présence de phényl isothiocyanate (PITC) qui réagit avec la fonction amine du résidu en position N-terminale pour former une liaison covalente. En présence d'acide trifluoroacétique, l'acide aminé en position N-terminale, toujours lié au PITC, est « décroché » de la chaîne polypeptidique pour former un dérivé thiazolinone. Ce dérivé est extrait dans un solvant organique et traité avec du H_3O^+ pour former un dérivé phénylthiohydantoïne qui est identifié par comparaison avec des standards.

IV.4.f. Expérience de thermofluorescence

Les expériences de thermofluorescence sont réalisées sur les échantillons de protéine soumis à des essais de cristallisation par la plateforme de l'EMBL HTXLab dirigée par José A. Marquez. Cette technique repose sur l'utilisation d'une sonde (comme le Sypro Orange) qui fluoresce lorsqu'elle se fixe de manière non spécifique sur les zones hydrophobes de la protéine. La dénaturation thermique de l'échantillon sous l'effet d'une augmentation de la température provoque l'exposition de nouvelles régions hydrophobes initialement enfouies et donc une fixation plus importante de la sonde, conduisant à un signal de fluorescence plus

important. Cette expérience apporte des informations sur les gammes de températures supportées par la protéine et sa température de fusion T_m .

IV.4.g. Expérience de dichroïsme circulaire

L'analyse des échantillons protéiques par dichroïsme circulaire a été réalisée à l'UVHCI sous la direction de Marc Jamin sur un spectropolarimètre 810 (JASCO), avec une vitesse de balayage de 1 nm/s en utilisant une cuve en quartz ($l=0,1$ nm) à une température de 20°C. Les échantillons contiennent de 3 à 20 μM de protéines dans **le tampon (NaP 20 mM, NaF 150 mM, pH 7,5)**.

Le spectre dichroïque du tampon est utilisé comme référence en début d'expérience et les valeurs obtenues sont ensuite soustraites des données recueillies pour les échantillons. Cinq spectres sont enregistrés pour chaque échantillon et sont moyennés pour augmenter la qualité du rapport signal / bruit. Les données sont enregistrées avec le logiciel Spectra manager (JASCO Corporation).

Cette technique repose sur l'utilisation d'une lumière polarisée (sélection de l'une des directions du champ magnétique, de manière soit linéaire, en choisissant une seule direction du champ magnétique mais à des amplitudes différentes, ou de manière circulaire en ne choisissant que les points du champ magnétique ayant les mêmes amplitudes mais dans des directions différentes).

En envoyant de la lumière polarisée circulairement à droite et à gauche sur un échantillon chiral celles-ci sont absorbées différemment. Dans une protéine, cette asymétrie est apportée par l'environnement de la liaison peptidique (principal chromophore aux longueurs d'ondes utilisées en dichroïsme circulaire 190-250 nm). L'asymétrie est due à l'orientation des différents plans peptidiques flexibles autour des liaisons N-C α , C α -CO et CO-N. Chacune de ces orientations est fonction des valeurs prises par les angles correspondants à ces liaisons, et chacun de ces angles aura une valeur caractéristique selon la structure secondaire.

Le spectrophotomètre de dichroïsme circulaire mesure la variation de cette absorbance. En décomposant la loi de Beer-Lambert entre la lumière droite et gauche elle devient :

$$\Delta A = A_g - A_d \text{ d'où } \Delta \epsilon = \epsilon_g - \epsilon_d = (\Delta A) / l.C$$

Comme la taille de la protéine influe sur la quantité de structure secondaire et donc l'absorbance, la valeur mesurée est ramenée au nombre d'acides aminés (n):

$$\Delta\varepsilon/n = (A_g - A_d)/l.c.n$$

Les mesures de l'absorbance dans le domaine des UV lointains (180-260 nm) sont caractéristiques de la composition en structures secondaires de la protéine :

- Les hélices α présentent un maximum d'absorbance à 195 nm et deux minima à 208 et 225 nm.
- Les feuillets β présentent à un maximum d'absorbance à 195 nm et un minimum à 220 nm
- Les structures non repliées (non-coiled) : présentent à un minimum d'absorbance à 200 nm et tendent vers une asymptote proche de 0 à partir de 215 nm.

IV.4.h. Spectroscopie de fluorescence

Les expériences de spectroscopie de fluorescence ont été réalisées à l'UVHCI sous la direction de Marc Jamin. Les échantillons protéiques sont à une concentration comprise entre 80 et 200 μM en **tampon Tris 20 mM, NaCl 150 mM, CaCl₂ 5 mM, pH 8**.

Certaines molécules (fluorophores) peuvent absorber des photons et se retrouver dans un état électronique excité. Le retour à l'état fondamental s'effectue par la libération rapide d'un photon de moindre énergie (sous forme d'émission de fluorescence). La majorité de la fluorescence intrinsèque d'une protéine provient de la présence des tryptophanes jouant le rôle de fluorophores. Dans l'eau le tryptophane absorbe la lumière à 280 nm et émet à 350 nm.

Cependant la longueur d'onde d'émission du tryptophane est dépendante de son environnement (notamment de la polarité de l'environnement). Cette émission a par exemple tendance à diminuer dans un environnement hydrophobe. En suivant l'émission de fluorescence des tryptophanes d'une protéine dans des conditions particulières il est possible de suivre l'évolution de sa conformation tridimensionnelle (exposition d'un tryptophane au solvant due à une dénaturation, atténuation de fluorescence due à une interaction avec un partenaire).

V. Méthodes d'analyse structurale des protéines étudiées

V.1. Détermination de structure atomique de protéine par cristallographie aux rayons X.

La cristallographie aux rayons X permet de déterminer la structure tridimensionnelle des molécules biologiques, comme l'ADN ou les protéines. Le principe de cette technique est très proche du mode de fonctionnement des microscopes pour lequel les rayons de la lumière visible sont diffusés par l'objet, puis focalisés grâce à plusieurs lentilles pour reconstituer une image de l'objet. Pour que le phénomène de diffusion soit possible, la longueur d'onde du rayon incident doit être du même ordre de grandeur que la taille de l'objet observé ($\lambda_{\text{lumière visible}}$ 400-800 nm pour des objets à observer de 0,1 à 1 mm). Pour résoudre la structure de macromolécules biologiques comme les protéines, nous souhaitons observer des distances interatomiques (de l'ordre de 1,5 Å) correspondant à la longueur des rayons X avec lesquels nous devons travailler. Cependant il n'existe aucune lentille capable de focaliser les rayons X pour reconstituer une image de la protéine. Le signal enregistré nécessitera donc un traitement informatique.

Les rayons X sont diffusés dans toutes les directions dans les protéines par les électrons gravitant autour des atomes. Mais les électrons étant de très mauvais diffuseurs, il est impossible d'enregistrer un signal suffisant à partir d'une molécule unique. La première étape en cristallographie des protéines aux rayons X est d'obtenir des cristaux de protéines pour amplifier le signal.

V.1.a. La cristallisation

V.1.a.1. Principe de la cristallisation des protéines

Un cristal est un empilement régulier et périodique d'un objet dans les trois dimensions de l'espace (un atome, une molécule). Sous un angle mathématique, un cristal est l'association d'un objet avec un réseau (**Figure 2.V.1**). Contrairement à un agrégat, dans un cristal les interactions entre protéines sont spécifiques.

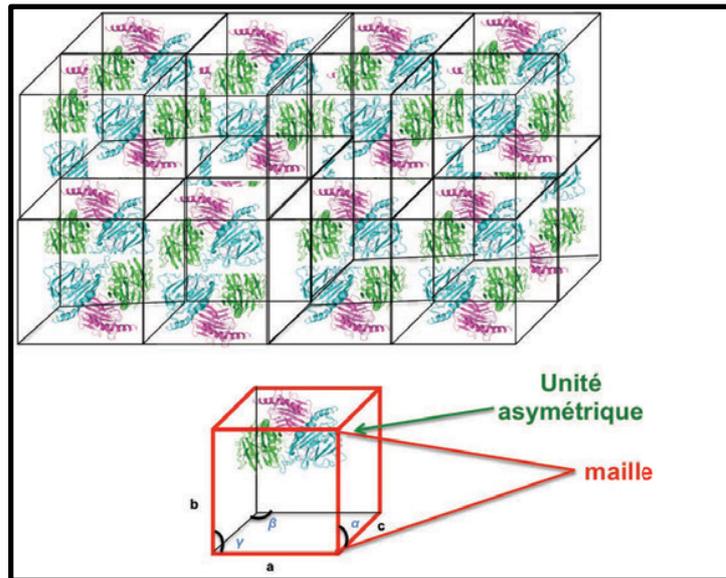


FIGURE 2.V.1 : REPRESENTATION SCHEMATIQUE D'UN ARRANGEMENT CRISTALLIN

Un cristal est un empilement périodique et régulier d'un objet dans l'espace, ici une protéine. Pour créer cet empilement les protéines interagissent entre elles à travers des contacts spécifiques (les mêmes résidus sont impliqués). La maille est l'unité la plus petite permettant de reconstruire l'ensemble du cristal par des opérations de translation. Elle peut contenir plusieurs copies de l'objet en question. Elle est définie par 3 paramètres de distance (a , b et c) ainsi que par 3 angles (α , β , γ). L'unité asymétrique est la plus petite unité constituant un cristal permettant de décrire le contenu de la maille par des opérations de symétrie (axes hélicoïdaux, rotoinversion, translation ...). L'objectif de la cristallographie des protéines aux rayons X est de résoudre la structure de l'unité asymétrique.

Dans un solvant, les protéines sont solubles grâce aux interactions qu'elles partagent avec les molécules d'eau. Pour les faire évoluer de l'état liquide à l'état solide (cristal), il est essentiel de diminuer leur solubilité en rompant les interactions protéine/solvant pour augmenter leur concentration et favoriser les interactions de type protéine/protéine. De nombreux paramètres physicochimiques peuvent influencer cette solubilité :

- La concentration initiale de la protéine, la présence d'impuretés, de ligands
- La température (pouvant influencer la cinétique des interactions)
- Le pH (influence la charge des résidus en surface et donc les possibilités d'interactions).
- La nature et la concentration de l'agent précipitant : les polymères comme le PEG qui excluent le solvant autour des protéines, les sels comme le sulfate d'ammonium qui entrent en compétition avec les liaisons ioniques protéines-solvant et forcent les interactions hydrophobes entre les protéines, et les solvants organiques qui diminuent la constante diélectrique du milieu.

L'influence de la concentration de la protéine en fonction des paramètres chimiques définit un diagramme de phase (**Figure 2.V.2**) pour une protéine donnée.

V.1.a.2. Technique de la goutte suspendue

La technique de la goutte suspendue ou de la diffusion en phase vapeur (**Wlodawer and Hodgson, 1975**) permet d'atteindre la limite de solubilité de la protéine en la concentrant. Une goutte de quelques μL de la solution protéique, mélangée à quelques μL de la solution de cristallisation est placée en suspension sur une lamelle de verre au dessus d'un puits contenant la solution de cristallisation (tampon, agent précipitant aux concentrations et pH prédéfinis). Le système est fermé hermétiquement (graisse ou joint selon les systèmes). La solution de cristallisation présente dans la goutte est alors diluée d'un certain facteur (selon le rapport de volumes goutte échantillon/goutte de solution de cristallisation) par rapport à celle contenue dans le puits. Au fil des jours, une mise en équilibre entre les deux solutions va s'établir par diffusion de vapeur, faisant évoluer la concentration de la goutte vers celle du puits. Cette équilibration se traduit par une concentration de la goutte et donc de l'échantillon protéique. Les conditions initiales sont choisies pour que le phénomène de concentration conduise l'échantillon protéique vers la zone de sursaturation, précisément de nucléation, essentielle pour provoquer la croissance cristalline (**Figure 2.V.2 et 2.V.3**).

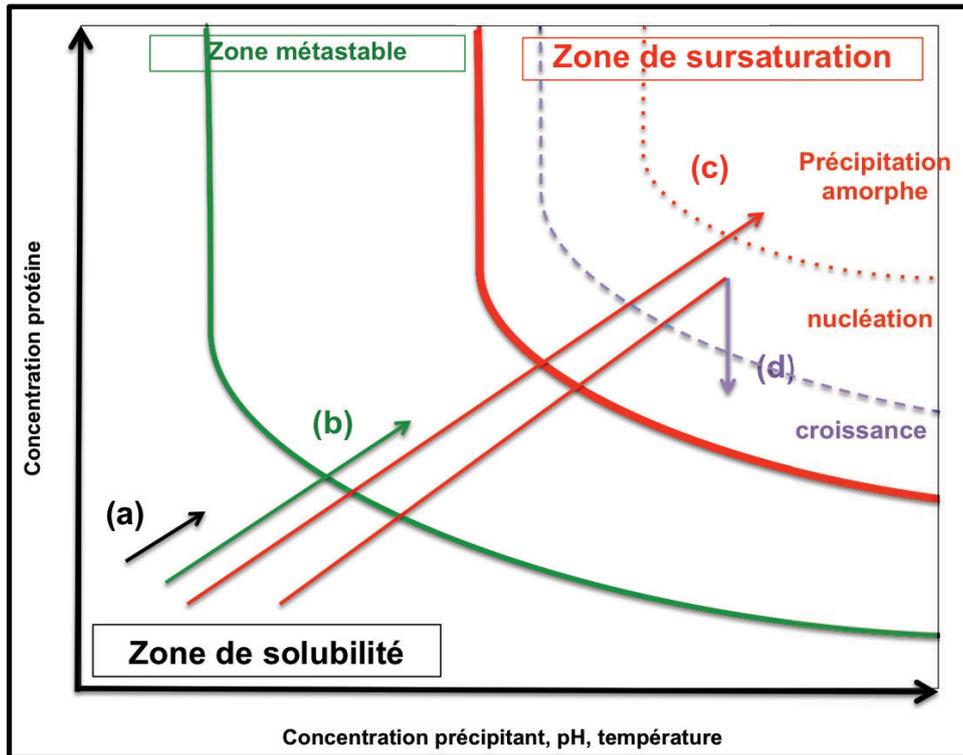


FIGURE 2.V.2 : DIAGRAMME DE PHASE D'UNE PROTEINE

Le diagramme de phase d'une protéine est constitué de 3 zones : une **zone de solubilité (a)** (la protéine est soluble, aucun cristal ne peut se former, un cristal introduit dans cette zone se dissout), une **zone métastable (b)** (aucun cristal ne peut se former, mais un cristal introduit dans cette zone reste stable, il ne grossit pas et ne fond pas), la **zone de sursaturation (c)** (divisée en deux parties : la **précipitation amorphe (d)** caractérisée par des interactions non spécifiques entre protéines et la présence d'agrégats de protéines, la **nucléation (e)** caractérisée par l'apparition de germes cristallins constitués d'interactions spécifiques et organisées entre protéines, (noyaux à partir desquels les cristaux vont grossir) et la **zone de croissance (f)** (dans laquelle les cristaux vont grossir à partir des noyaux/ germes)).

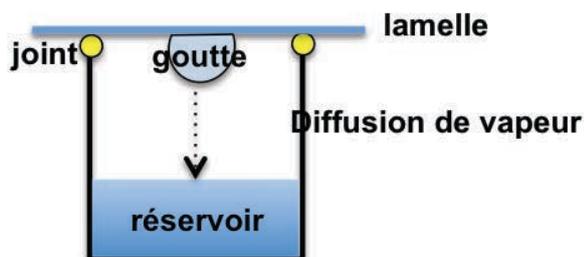


FIGURE 2.V.3 : TECHNIQUE DE LA GOUTTE SUSPENDUE

L'échange de vapeur entre le réservoir et la goutte en suspension permet de concentrer la protéine présente dans la goutte. Les conditions sont choisies pour atteindre la zone de nucléation du diagramme de phase. Une fois dans cette zone, des germes vont apparaître. La concentration de la protéine en solution diminue, le système atteint alors la zone de croissance dans laquelle les cristaux vont grossir (**Figure 2.V.2 d et f**).

V.1.b. Congélation des cristaux de protéines

Les rayons X de hautes intensités peuvent générer d'importants dommages sur les protéines lors de l'enregistrement des clichés de diffraction à travers la formation de radicaux libres et donc nuire à la qualité de la diffraction. Pour réduire la propagation des dommages radiatifs, il faut avoir recours à la cryocristallographie (congélation du cristal dans de l'azote liquide pour atteindre une température inférieure à 100°K). Les cristaux contiennent cependant de grandes quantités de solvant aqueux (30 à 80%) et les congeler dans ces conditions entraînerait la cristallisation de l'eau qui peut abîmer le cristal (contrainte mécanique), mais également le cliché de diffraction (bruit de fond apporté par la diffraction de la glace). Pour éviter cela, un cryoprotectant est utilisé (PEG, glucose, glycérol...). Ce choix est important, car l'ajout de certains cryoprotectants pour un cristal donné peut l'endommager.

Les conditions ayant permis la formation des cristaux utilisés pour la résolution des structures des fragments de calréticuline de *T. cruzi* et *E. histolytica* ainsi que les cryoprotectants utilisés sont répertoriés dans les **tableaux 2.V.1 et 2.V.2**

| Protéine cristallisée | Concentration échantillon | Ratio gouttes (protéine/tampon/additif) (μ L) | Conditions de cristallisation | Solution de cryo protection |
|----------------------------------|---------------------------|--|---|--|
| TcCRT Δ Pter Δ 288 | 12 mg/mL | 1/1 | 35% PEG 4000 0,2 M Sodium acetate 0,1 M Tris pH 8,5 | Tampon de cristallisation + 10% glycérol |
| | | 1/1/1 (glucose 1 M) | 40% PEG 4000 0,2 M sodium acetate 333 mM glucose 0,1 M Tris pH 8,5 | aucun |

TABLEAU 2.V.1: CONDITIONS DE CRISTALLISATION DU FRAGMENT TcCRT Δ Pter Δ 288

| Protéine cristallisée | Concentration échantillon | Ratio gouttes (protéine/tampon/additif) (μL) | Conditions de cristallisation | Solution de cryo protection |
|-------------------------|---------------------------|---|---|--|
| EhCRTAPter Δ 288 | 11 mg/mL | 1/1 | 2,7 M ammonium sulfate 0,1M Tri-Sodium citrate pH 5 | 3 M ammonium sulfate 0,1 M Tri-Sodium citrate pH 5 25% glycérol |
| | | 1/1/1 (Phosphosérine 20 mM) | 2,6 M ammonium sulfate 7 mM phosphosérine 0,1 M Tri-sodium citrate pH 5 | 3 M ammonium sulfate 0,1 M Tri-Sodium citrate pH 5 20 mM phosphosérine 25% glycérol |

TABLEAU 2.V.2: CONDITIONS DE CRISTALLISATION DU FRAGMENT EHCRTAPTER Δ 288

V.1.c. Enregistrement des clichés de diffraction

Comme présenté ci-dessus, les électrons présents dans les protéines sont responsables de la diffusion des rayons X. Les rayons X sont des ondes (avec une composante magnétique et une composante électrique). Etant donné le grand nombre d'électrons présents, les ondes diffusées peuvent interférer, c'est le phénomène de diffraction. Du fait de l'utilisation d'un cristal, ce phénomène de diffraction est discret, c'est-à-dire que les rayons X ne sont diffusés que dans certaines directions du fait des interférences constructives et destructives des ondes. Les interférences constructives sont responsables de l'amplification du signal et il est possible de déterminer les directions dans lesquelles les ondes diffusées sont en phase (interférence constructive) en utilisant la loi de Bragg (**Figure 2.V.4**).

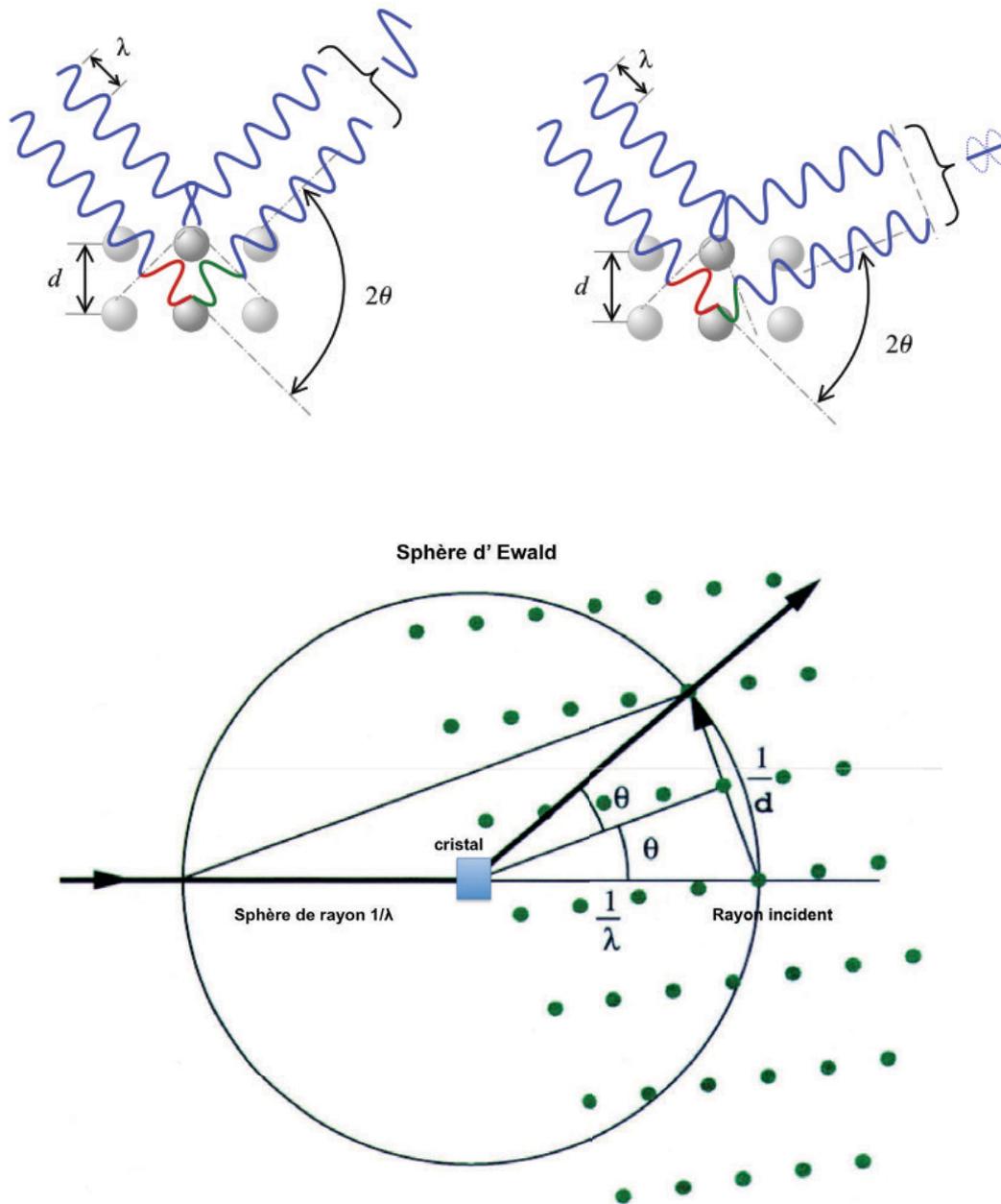


FIGURE 2.V.4: SCHEMAS DECRIVANT LE PHENOMENE DE DIFFRACTION SELON LA LOI DE BRAGG

Un cristal étant la répétition d'un même objet, il est possible de le diviser en des plans contenant tous les mêmes atomes. Ces plans réticulaires sont définis par les indices h,k,l . Selon la loi de Bragg, les ondes diffusées sont en phase (et il y a donc interférence constructive) si l'équation $n\lambda = 2d\sin\theta$ est vérifiée (avec λ la longueur d'onde du rayon incident et n un entier, d la distance entre deux plans et θ l'angle entre le faisceau incident et le plan). Le réseau réciproque est un réseau 3D dans l'espace de diffraction, chaque tache de diffraction d'indice hkl , correspondant à l'interférence constructive de tous les plans du cristal d'indice hkl . De ce fait la loi de Bragg peut également être représentée graphiquement par la sphère d'Ewald. Lorsqu'un point du réseau réciproque traverse la sphère d'Ewald la loi de Bragg est vérifiée et il y a diffraction.

V.1.c.1. Le synchrotron

Les synchrotrons sont de grands instruments permettant de générer des rayons X de très haute intensité (5 à 25 keV). Les électrons sont accélérés par paquets sous vide (10^{-13} atm pour éviter une perte de vitesse et donc d'énergie par collision avec les molécules de l'air) à une vitesse proche de celle de la lumière pour atteindre une énergie proche de 2 GeV dans un accélérateur linéaire puis un anneau d'accélération grâce à des accélérateurs circulaires. Les paquets d'électrons sont ensuite transférés dans un anneau de stockage, dans lequel des aimants de courbure (envoyant des impulsions magnétiques de 2 Tesla perpendiculaires à la trajectoire des électrons) imposent une trajectoire aux électrons. Ces aimants leur font subir une accélération qui se traduit par l'émission d'un rayonnement : le rayonnement synchrotron (faisceau de photons s'étendant de l'infrarouge aux rayons X durs, de manière tangentielle à la trajectoire des électrons).

Les rayons X générés sont canalisés dans les lignes de lumière qui contiennent des onduleurs et des wigglers (aimants fournissant un champ magnétique alternatif) permettant d'augmenter l'intensité du rayonnement. Il est alors possible de sélectionner la longueur d'onde du faisceau polychromatique incident, qui est peu divergent et polarisé. La puissance des synchrotrons actuels permet de réduire considérablement la durée de l'expérience : pour un cliché de diffraction, l'exposition du cristal se situe entre 0,01 s et quelques minutes.

V.1.c.2. Choix de la stratégie de diffraction

Les différents jeux de données utilisés dans ma thèse ont été obtenus à l'ESRF de Grenoble sur les lignes de lumière ID29 et ID23. Comme je l'ai présenté dans la section précédente et **figure 2.V.4**, les ondes diffusées n'interagissent de manière constructive que dans certaines directions. Pour obtenir la totalité de l'information permettant de résoudre la structure de la protéine, nous devons effectuer une rotation du cristal pour que tous les noeuds du réseau réciproque traversent la sphère d'Ewald et ainsi enregistrer la totalité des taches de diffraction dans toutes les directions de l'espace. Cependant, selon la loi de Friedel, le réseau réciproque (donc le cliché de diffraction) est centro-symétrique, c'est-à-dire que l'intensité correspondant à la diffraction des plans h, k, l (I_{hkl}) est la même que l'intensité correspondant à la diffraction des plans $-h, -k, -l$ (I_{-h-k-l}), ce qui signifie que même en absence de symétrie dans le cristal il n'est nécessaire d'enregistrer que 180° pour obtenir un jeu de données complet. Plus la symétrie dans le cristal est importante, plus nous enregistrons des taches de diffraction équivalentes et donc moins il est nécessaire d'enregistrer d'images pour obtenir un jeu complet (45° seraient suffisant pour des axes de symétries d'ordre 4). Bien entendu, il est

toujours intéressant, selon la résistance du cristal, d'enregistrer un maximum de données pour améliorer la redondance et la précision des données.

Définir la stratégie de la collecte des données nécessite donc de choisir le nombre de clichés de diffraction à enregistrer et à quelle fréquence (par exemple il est possible d'enregistrer un cliché tous les 1° ou $0,1^\circ$ et dans ce dernier cas il faut 1800 images pour couvrir 180°), le temps d'exposition du cristal aux rayons X pour chaque orientation ainsi que la puissance du faisceau (transmission). Ces dernières doivent être assez grandes pour enregistrer l'information de plus faible intensité, et assez faibles pour que le cristal supporte les dommages radiatifs le temps de l'expérience. Les détecteurs PILATUS installés sur les lignes ID23 et ID29, permettent de travailler avec des temps d'exposition très courts, améliorant la survie du cristal et permettant de s'affranchir du bruit de fond de manière importante.

A ce stade il est également essentiel de centrer le cristal et si possible aligner le faisceau perpendiculairement à l'axe de la maille ayant le plus haut niveau de symétrie.

Le logiciel **EDNA** (Incardona et al., 2009) a largement été utilisé au cours de cette thèse pour nous aider à optimiser la stratégie d'enregistrement des données de diffraction, en particulier avec les nouveaux détecteurs Pilatus.

V.1.d. Traitement des clichés de diffraction

V.1.d.1. XDS : Indexation, Intégration, Mise à l'échelle

Le traitement des jeux de données obtenus au synchrotron se divise en différentes étapes. Pour les réaliser, le programme XDS a été utilisé (XRays data Streamer (Kabsch, 2010)). Ce programme est composé de sept sous-logiciels qui s'enchaînent automatiquement pour réaliser les différentes étapes.

- Le logiciel **XYCORR**, **INIT**, permet de réduire le bruit de fond en calculant les corrections de distorsions spatiales à partir d'une image de calibration.
- **COLSPOT** établit une liste des coordonnées des réflexions les plus intenses.
- **IDXREF** utilise le fichier de sortie de COLSPOT et indexe les réflexions pour déterminer les paramètres de maille (a , b , c , α , β , γ) et nous propose des groupes d'espaces potentiels en comparant le profil obtenu avec une banque de groupes d'espaces et de paramètres de maille. Cela permet d'attribuer les indices h,k,l à chaque tache de diffraction
- **COLPROF** établit le profil tridimensionnel des taches de diffraction, afin de pouvoir mesurer précisément leur intensité à l'étape suivante. Cela permet de gérer les cas de

mosaïcité et les stratégies utilisant de petits angles de rotation du cristal, pour lesquels l'intensité d'une tache peut se répartir sur plusieurs images.

- **INTEGRATE** intègre l'intensité de chaque tache de réflexion à partir du profil calculé. Permettra de calculer les facteurs de structures (h,k,l).
- **CORRECT, XSCALE** permet de mettre à l'échelle les réflexions équivalentes présentes sur plusieurs images. L'intensité d'une même réflexion sur plusieurs images doit être théoriquement la même, mais ce n'est jamais le cas à cause des différences d'absorption et de la décroissance de la diffraction du cristal.
- **GLOREF/CELL PARM** affine les paramètres de maille.

A la fin de ce traitement, nous obtenons un fichier (ASCII) permettant de générer un fichier mtz contenant les informations sur le cristal, et les amplitudes des facteurs de structure.

V.1.d.2. Qualité des données

A la fin du traitement des données, plusieurs paramètres permettent d'évaluer la qualité de ces dernières.

- **La résolution** : détermine la finesse de l'image, c'est la distance minimale (exprimée en Å) pour laquelle nous sommes capables de séparer deux objets très proches. Elle peut être limitée par l'expérience (taille du détecteur, distance cristal-détecteur) ou par la qualité de diffraction du cristal. Lors de cette thèse il a été choisi de garder la tranche de résolution pour laquelle le coefficient de corrélation ($CC_{1/2}$) est au moins de 80% et le rapport $I/\sigma(I)$ supérieur à 2.
- **Le rapport signal/bruit ($I/\sigma(I)$)** : permet d'avoir une estimation de l'intensité moyenne des réflexions mesurées à une résolution par rapport au bruit de fond.
- **La complétude** : indique le nombre de taches de diffraction mesurées par rapport au nombre de taches théoriquement possibles en fonction du groupe d'espace sélectionné (doit être supérieur à 90%).
- **La redondance** indique le nombre moyen de répétitions de la mesure d'une réflexion et de ses symétriques. Plus grande elle est, meilleure sera l'estimation de l'intensité de la réflexion.
- **Le facteur R_{sym} ($\sum_{\text{hkl}} |I_{\text{hkl}} - I| / \sum_{\text{hkl}} I_{\text{hkl}}$)** : compare l'intensité de chaque réflexion équivalente par la symétrie du cristal (I_{hkl}) à la valeur moyenne de ces réflexions (I). Plus la valeur est basse, meilleure est la cohérence du jeu de données.
- **$CC_{1/2}$** : Le facteur de corrélation $1/2$ est une estimation expérimentale de la valeur finale des données fusionnées avec l'idée sous-jacente que les valeurs soient réelles. C'est le

coefficient de corrélation de Pearson de deux demi-jeux de données (chacun est dérivé de la moyenne de la moitié des observations pour une réflexion donnée). Il est actuellement plus fréquemment utilisé que les critères $I/\sigma I$ et le R_{sym} pour juger de la qualité des données et est aujourd'hui décrit comme le meilleur critère (**Diederichs and Karplus, 2013**).

V.1.d.3. Paramètres du cristal

Lors du traitement des jeux de données plusieurs paramètres liés au cristal sont importants et doivent être déterminés.

- **Le groupe d'espace** : Comme je l'ai expliqué, la maille permet de construire l'empilement cristallin par des opérations de translation. Il n'existe qu'un nombre limité de structures géométriques permettant de remplir l'espace sans laisser de vide en effectuant des translations, ce qui détermine les sept systèmes cristallins. Pour respecter la symétrie au sein du cristal en n'effectuant que des translations, l'objet dans le cristal ne peut être présent qu'à certaines positions pour chaque système cristallin (ce qui détermine les 14 réseaux de Bravais). En associant à chaque réseau de Bravais les opérations de symétrie présentes dans la maille pour la « reconstruire » à partir de l'unité asymétrique, nous définissons 64 groupes d'espaces possibles (parmi 230 possibles, en effet les symétries faisant intervenir une inversion de la chiralité des molécules, comme les symétries miroir ne sont pas autorisées pour les protéines) (**Figure 2.V.5 et Tableau 2.V.3**). En fonction des paramètres de maille et des réflexions équivalentes, il est nécessaire de déterminer le bon groupe d'espace pour être capable de calculer la position correcte de la molécule dans la maille.
- **La mosaïcité du cristal** : La mosaïcité caractérise l'empilement cristallin, pouvant être influencé lors de la croissance cristalline par des contraintes mécaniques et thermiques, pouvant également être apporté par la congélation. Cette mosaïcité peut entraîner des problèmes de superposition des taches de diffraction.
- **Le facteur d'agitation thermique (B)** : Les protéines présentes dans le cristal peuvent avoir des zones flexibles et/ou agitées dépendant de la température et de la compacité du cristal. Cette agitation entraîne un affaiblissement global des intensités diffractées, ce qui est pris en compte par le facteur d'agitation thermique global B (exprimé en Å^2).

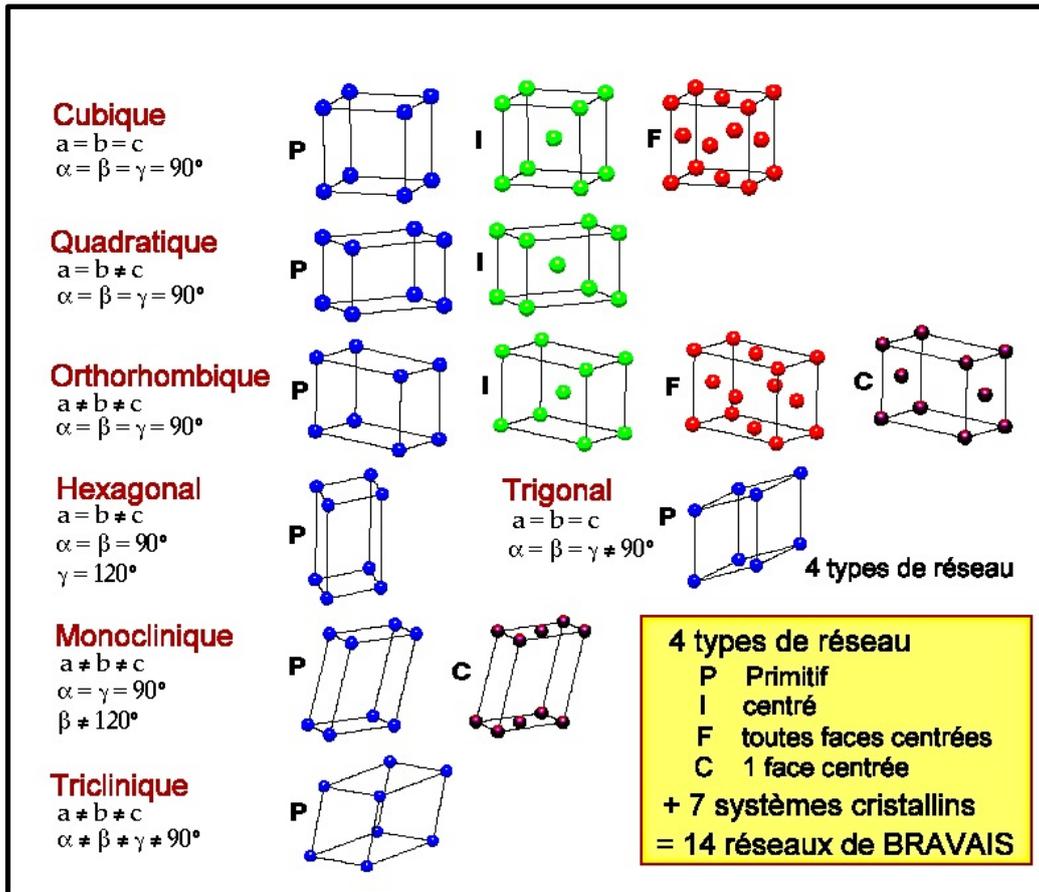


FIGURE 2.V.5: DESCRIPTION DES SYSTEMES CRISTALLINS ET RESEAUX DE BRAVAIS

| Systèmes cristallins | Exemple de groupes d'espaces |
|----------------------|---|
| Triclinique | P1 |
| Monoclinique | P2, P2 ₁ , C2 |
| Orthorhombique | P222, P2 ₁ 2 ₁ 2 ₁ , P2 ₁ 2 ₁ 2... |
| Quadratique | P4, P422, I4 ₁ 22... |
| Rhomboédrique | P3, P312... |
| Hexagonal | P6, P6 ₂ |
| Cubique | P23, P4 ₁ 32, I432... |

TABLEAUX 2.V.3: SYSTEMES CRISTALLINS ET EXEMPLES DE GROUPES D'ESPACE ASSOCIES

Les systèmes cristallins sont classés en fonction de leur degré de symétrie. Des exemples de groupes d'espace associés à chaque système sont présentés. Par exemple le système triclinique ne présente aucune symétrie, alors que le groupe d'espace P222 en orthorhombique présente trois axes de symétrie d'ordre 2 le long des 3 axes a, b et c. Le groupe dérivé P2₁2₁2₁ possède en plus 3 axes hélicoïdaux (rotation de 180° suivie d'une translation de 1/2a le long de l'axe a par exemple).

V.1.e. Résolution de la structure cristallographique

Au cours de cette thèse, les différentes étapes de la résolution de structures ont été effectuées avec les logiciels proposés par la suite CCP4 (coefficient de Matthews, détermination de la phase, calcul de carte de différence anormale, affinement) ou le logiciel Phénix (affinement, validation de la structure).

V.1.e.1. Le problème de la phase

Comme je l'ai déjà mentionné, les rayons X sont une onde. L'interférence des ondes diffusées par le cristal peut être représentée par un vecteur : le facteur de structure (**Figure 2.V.6**).

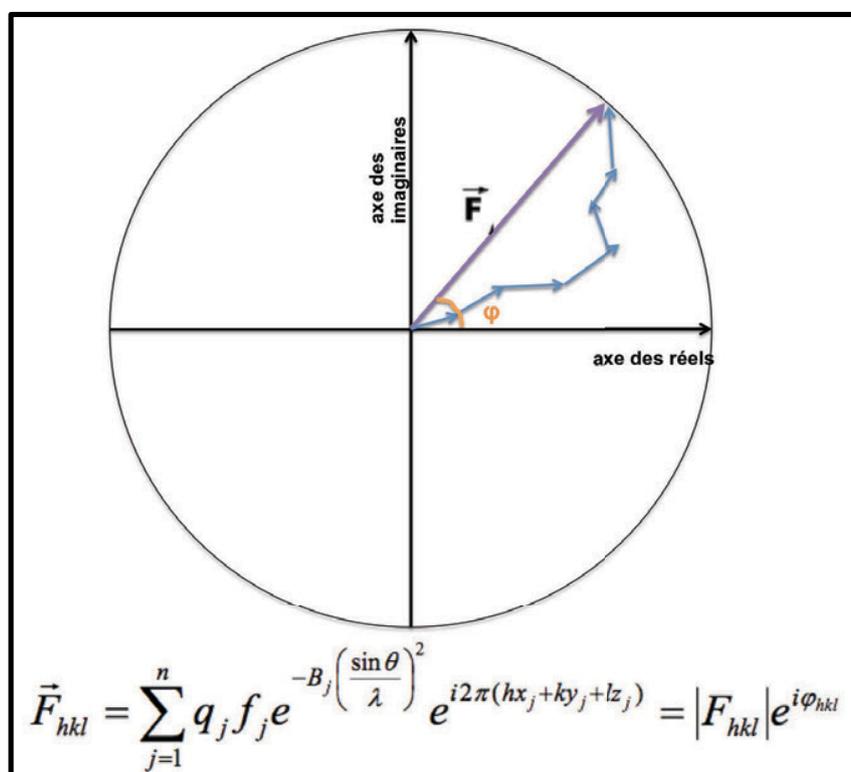


FIGURE 2.V.6: LE FACTEUR DE STRUCTURE

Le facteur de structure est la somme de tous les facteurs de diffusion des atomes présents dans le cristal (vecteurs bleus) et est caractérisé par une amplitude (longueur du vecteur) et une phase (ϕ). Il contient une composante réelle et une composante imaginaire. En cristallographie, nous n'avons pas la résolution pour observer les orbitales atomiques, de ce fait nous raisonnons avec une densité électronique sphérique et centro-symétrique et le facteur de structure peut être exprimé mathématiquement grâce à une série de Fourier (avec q_j , B_j , x_j , y_j , z_j : l'occupation, le facteur d'agitation et les coordonnées spatiales de l'atome j).

Grâce à la transformée de Fourier inverse, nous pouvons recalculer l'image de la densité électronique. Il est donc très important de décrire le facteur de structure en amplitude et en phase pour calculer la densité électronique.

$$\rho(\vec{r}) = \frac{1}{V} \sum_h \sum_k \sum_l \vec{F}_{hkl} e^{-i2\pi(hx_j + ky_j + lz_j)}$$

Cependant lorsque nous enregistrons des données de diffraction, la seule information que nous obtenons hormis les paramètres de maille est l'intensité de chaque tache de diffraction. Cette intensité est directement reliée et proportionnelle à l'amplitude du facteur de structure.

$$I_{hkl} = \frac{\lambda^3}{\omega \cdot V^2} \times \left(\frac{e^2}{mc^2} \right)^2 \times V_{cr} \times I_0 \times L \times P \times T_r \times |\vec{F}_{hkl}|^2 \propto |\vec{F}_{hkl}|^2$$

Nous perdons donc une information essentielle au calcul de la densité électronique : la phase.

V.1.e.2. Estimer le nombre de molécules dans l'unité asymétrique : le coefficient de Matthews

Pour déterminer correctement les phases, il est tout d'abord essentiel d'estimer le nombre de molécules de protéines contenues dans l'unité asymétrique qui peuvent être reliées entre elles par des symétries non cristallographiques. Pour cela nous utilisons le coefficient de Matthews (**Kantardjieff and Rupp, 2003; Matthews, 1968**). Matthews a démontré que dans la majorité des cristaux de protéines le rapport (Vm) entre le volume de la maille (V) et la masse moléculaire (M) contenue dans celle-ci était compris entre 2 et 4 Å³/Da en utilisant la relation $V_m = (V / M.n.Z)$, avec n le nombre de positions équivalentes de la protéine dans l'unité asymétrique et Z le nombre de molécules dans l'unité asymétrique. De la même manière, le pourcentage de solvant (S) peut être déterminé par la relation $S = 1 - (1,23/V_m)$ (1,23 g/cm³ est la valeur de la masse volumique moyenne des protéines).

Cette méthode nous permet donc d'estimer la quantité de solvant présent dans le cristal, ainsi que le nombre de protéines dans l'unité asymétrique. Même si le coefficient de Matthews donne des bons scores pour des pourcentages de solvant situés entre 30 et 50%, il n'est pas rare aujourd'hui, avec l'amélioration des lignes de lumières (puissance, détecteurs..), d'enregistrer des jeux de données sur des cristaux contenant de 60 à 80% de solvant.

V.1.e.3. Le remplacement moléculaire : Phaser

Il est admis que si deux protéines possèdent une identité de séquence supérieure à 30%, il existe de fortes probabilités qu'elles présentent un repliement tridimensionnel similaire. La

technique du remplacement moléculaire introduit en 1962 (**Rossmann, 2001**) repose sur ce principe. Il s'agit d'utiliser le modèle similaire pour calculer une densité électronique de départ (en utilisant les intensités expérimentales et les phases du modèle). Pour cela le processus est décomposé en deux étapes :

- Une étape de rotation, pour définir l'orientation du modèle dans la maille.
- Une étape de translation, pour positionner correctement le modèle dans la maille.

Dans ces deux étapes, le remplacement moléculaire utilisait la fonction de Patterson, une fonction d'autocorrélation considérée comme la transformée de Fourier des intensités diffractées. Elle présente des maxima pour les valeurs correspondant aux distances interatomiques. De ce fait elle représente les vecteurs entre atomes d'une même molécule ou de molécules distinctes reliées par symétrie.

Pour trouver l'orientation du modèle, la méthode consiste à déterminer les angles d'Euler ($\theta_1, \theta_2, \theta_3$) pour lesquelles les fonctions de Patterson du modèle orienté et celle issue du jeu de données sont superposées au maximum. Une fois correctement orienté en rotation, le modèle est placé dans la maille par translation en utilisant les vecteurs intermoléculaires (entre molécules symétriques) de la fonction de Patterson.

Dans le cadre de cette thèse, la totalité des structures ont été résolues en utilisant la méthode du remplacement moléculaire avec le programme Phaser de la suite CCP4 (**McCoy et al., 2007**). Ce logiciel se base sur la technique du maximum de vraisemblance pour déterminer les rotations et translations à appliquer au modèle. L'accord entre le modèle et les données expérimentales est donné par un Z-score (s'il est inférieur à 5 aucune solution n'a été trouvée, entre 5 et 6 la solution est incertaine, entre 6 et 7 elle est possible, entre 7 et 8 elle est probable et supérieure à 8 la solution est définitive). Pour utiliser cette technique il peut être nécessaire de supprimer du modèle utilisé certaines zones flexibles ou identifiées comme trop éloignées de la protéine d'intérêt.

V.1.e.4. Diffusion anormale

J'ai précédemment introduit que la diffusion des rayons X par les électrons était une diffusion cohérente, c'est à dire que la longueur d'onde et l'énergie des faisceaux incidents et diffusés sont les mêmes. Toutefois dans des cas particuliers, en sélectionnant une longueur d'onde précise pour que l'énergie des rayons X corresponde au seuil d'absorption d'un atome, les photons du rayon incident entrent en résonance avec les électrons de l'atome et le faisceau

diffusé est alors déphasé par rapport au faisceau incident. C'est le phénomène de diffusion anormale. Le facteur de diffusion de l'atome s'incrémente d'une valeur imaginaire correspondant à ce phénomène de résonance et ce dernier a pour effet de briser la loi de Friedel, c'est-à-dire que les intensités correspondant à la diffraction des plans hkl sont cette fois différentes de celles des plans -h-k-l. Dans ce cas de figure précis, il est alors nécessaire d'enregistrer le signal sur 360°.

En introduisant un atome lourd ou un diffuseur anormal à la protéine (ou si ceux-là sont déjà présents dans la protéine ou le tampon de cristallisation et organisés), nous pourrions identifier la position de cet atome et calculer des phases expérimentales. Dans le cadre de cette thèse, la diffusion anormale a été utilisée pour confirmer la présence de certains atomes (S, Cl).

V.1.e.5. Calcul des cartes de densité

L'estimation des phases, quelles soient expérimentales ou issues d'un modèle homologue, permet de calculer les facteurs de structure et donc des cartes de densité électronique. Il existe différents types de carte de densité :

- Les cartes de différence : Les cartes de différence $2F_{\text{obs}}-F_{\text{calc}}$ et $F_{\text{obs}}-F_{\text{calc}}$ rendent compte des erreurs entre les facteurs de structures observés et ceux calculés pouvant contenir des erreurs. C'est notamment le cas en remplacement moléculaire où la densité calculée est souvent biaisée par les phases apportées par le modèle. La première nous permet de reconstruire le modèle en montrant la densité pour les erreurs du modèle, tandis que la seconde pointe les erreurs du modèle à travers des pics positifs (indiquant une zone où il manque des électrons) et des pics négatifs (zone présentant un excès d'électrons, dans le cas où des atomes ont été placés alors qu'il n'y a aucune information expérimentale dans cette zone).
- Les cartes omit : Ces cartes sont utilisées pour confirmer la nature d'une densité électronique en s'assurant qu'elle n'est pas due à un biais apporté par le modèle. Pour cela, nous calculons une carte de densité en supprimant la partie du modèle en question, les phases de cette zone devant être apportées par le reste du modèle. Ceci permet de confirmer la présence réelle d'un ligand.
- Les cartes de différences anormales : Les cartes de différences anormales sont calculées à partir des réflexions de jeux de données enregistrés sur 360° à la longueur d'onde assez éloignée dans notre cas du seuil d'absorbance du diffuseur anormal (dans notre cas le chlore et le soufre) mais suffisamment proche pour qu'il reste un signal

suffisant. La carte de densité est calculée après soustraction des facteurs de structure F^+ et F^- (correspondant aux paires de Friedel). Elle met en évidence des pics positifs correspondant au nuage électronique des diffuseurs anomaux.

V.1.f. Amélioration du modèle et des cartes de densité

V.1.f.1. Modification manuelle et affinement

Comme je viens de l'introduire, les cartes de densité calculées au départ comportent souvent des erreurs, notamment dans le cas du remplacement moléculaire où elles contiennent un biais apporté par les phases du modèle. Plusieurs alternatives sont possibles pour améliorer le modèle et les cartes de densités.

- La première stratégie est la reconstruction manuelle, effectuée en utilisant le programme Coot. Cette étape est nécessaire car les programmes d'affinement numérique ne modifient que légèrement la position des atomes sans changer le modèle. Par exemple une étape clé, lorsque l'on travaille avec un modèle issu du remplacement moléculaire, est la substitution des acides aminés pour travailler avec la séquence de la protéine d'intérêt (cette étape permet de vérifier la qualité des cartes, par exemple la présence de pics positifs à un endroit où la protéine d'intérêt contient une insertion par rapport au modèle). Les modifications manuelles sont également nécessaires pour corriger les rotamères des chaînes latérales, les différents angles etc...
- L'affinement numérique permet de compenser les erreurs apportées par la reconstruction manuelle, et minimise l'écart entre les facteurs de structures observés et calculés en s'appuyant sur la méthode des maximums de vraisemblance, ou la dynamique moléculaire.

V.1.f.2. Refmac/ Phenix Refine

Pour effectuer ces affinements numériques, deux programmes ont été largement utilisés durant cette thèse : Refmac5 de la suite CCP4 (**Murshudov et al., 1997**) et Phenix Refinement, un sous-programme de Phenix (**Adams et al., 2010**).

Ces deux programmes se basent sur une méthode de minimisation prenant en compte les valeurs des facteurs calculées à partir du modèle et les valeurs observées et essaie de minimiser une fonction correspondant à l'énergie potentielle du modèle (ces différents

facteurs correspondent aux liaisons, angles, coordonnées spatiales des atomes, agitation thermique, énergie de van der Waals):

- Le maximum de vraisemblance : estime la vraisemblance des facteurs de structure du modèle en estimant la probabilité de calculer les F_{obs} si le modèle est juste. Cette fonction de probabilité utilise les données exclues de l'affinement (R_{free} et permet d'atteindre plus facilement le minimum des facteurs R et R_{free}).

Le logiciel Refmac5 détermine des contraintes géométriques à partir du modèle, puis calcule les facteurs de structure du modèle et les met à l'échelle par rapport à ceux observés en utilisant les méthodes de minimisation.

Le programme Phénix Refinement travaille globalement de la même manière. Il commence le cycle d'affinement par un aplatissement de solvant et travaille avec deux fonctions d'énergie différentes à minimiser. La première est basée sur la densité électronique (real space refinement) et la seconde sur les facteurs de structure (reciprocal space refinement). Il offre également une possibilité d'affinement en utilisant le recuit simulé (dynamique moléculaire).

Ces programmes nous offrent le choix des paramètres que nous souhaitons incrémenter aux différentes fonctions d'affinement (facteur B, coordonnées x,y,z, occupations, angles, liaisons, choix du poids de la géométrie du modèle/densité électronique). Phenix propose également d'affiner en fonction d'un modèle référence à haute résolution ou de structure secondaire. Parmi ces paramètres, deux sont très couramment utilisés et augmentent grandement la qualité du modèle :

- **NCS (Non crystallographic symmetry)** : Dans le cas où dans l'unité asymétrique des molécules sont reliées entre elles par des symétries non cristallographiques, ce paramètre permet de remplacer la densité électronique correspondant à chaque région concernée par la densité moyenne. Cela apporte des contraintes dans les facteurs de structure.
- **TLS (Translation/Libration/Screw)** : Il s'agit d'un modèle mathématique prédisant le déplacement local des atomes en se basant sur l'hypothèse que chaque atome se déplace comme un corps rigide par rapport à une position moyenne. Ce modèle contient les contributions provenant des vibrations de la molécule dans le cristal, ce qui apporte de l'information sur les micro-conformations de groupes d'atomes dans la

molécule. Les atomes d'un même bloc regroupés selon leurs agitations sont affinés ensemble par rapport à une position moyenne.

Le nombre de réflexions enregistrées prend ici toute son importance. En effet plus le rapport observations/paramètres à affiner est important, meilleur sera l'affinement et la qualité des cartes et du modèle. Par exemple pour un jeu de données à basse résolution (peu d'observation) nous choisirons un affinement des facteurs B de manière isotropique, alors que pour une haute résolution nous serons capables d'affiner les facteurs B dans les trois directions x, y et z.

V.1.f.3. Validation du modèle

Au cours des cycles affinement/reconstruction, de nombreux paramètres nous permettent d'évaluer la qualité du modèle et des cartes de densité.

$$R = \frac{\sum_{hkl} \left| |F_{obs}| - |F_{calc}| \right|}{\sum_{hkl} |F_{obs}|}$$

- Rfact : permet d'estimer l'accord entre les facteurs de structure calculés et ceux observés. Il doit être le plus bas possible.
- Rfree : se calcule de la même manière que le Rfact mais sur un ensemble représentant 5 à 10% des réflexions qui ont été exclues de l'affinement (**Brünger, 1992**). En effet, il peut arriver que le modèle biaise le calcul des phases et donc des cartes. Le facteur Rfree doit donc évoluer comme le Rfact: ils doivent diminuer de concert et être le plus proche possible. C'est un bon repère pour estimer la qualité des phases calculées. En cas de sur affinement du modèle, ce facteur aura tendance à augmenter (le facteur Rfact diminuera plus vite que le Rfree).
- Les différents paramètres géométriques : Le modèle doit respecter les distances et les angles conventionnels interatomiques, la géométrie des chaînes latérales.
- Le diagramme de Ramachandran : Les angles Φ et Ψ de la liaison peptidique situés en amont et en aval de la chaîne latérale de chaque acide aminé, ne peuvent prendre qu'un nombre de valeurs déterminées à cause de l'encombrement stérique apporté par la chaîne latérale. Pour chaque résidu, ces angles « autorisés » représentant les régions énergétiquement favorables sont représentés sur le diagramme de Ramachandran qui doit être respecté.
- Le clashscore : Permet d'évaluer le recouvrement (overlap) des nuages électroniques de deux atomes situés dans un environnement proche. Pour des raisons énergétiques

(répulsion des nuages électroniques) et d'encombrement stérique, la distance entre deux atomes donnés ne peut être inférieure à une certaine valeur, correspondant à la somme de leurs rayons de van der Waals.

Le programme Molprobity permet d'évaluer ces différents critères en nous imposant des limites. Il nous permet également de comparer toutes ces valeurs avec l'ensemble des structures déposées dans la Protein Data Bank, résolues à la même résolution. Cela permet d'évaluer la qualité du modèle. Ce programme étant lié à Phenix Refinement durant l'affinement, ce dernier a tendance à mettre un poids sur la géométrie plus important que Refmac5 au détriment des facteurs de structures. Au cours de cette thèse le programme Refmac5 a souvent été utilisé pour améliorer rapidement les cartes, tandis que Phenix Refinement a été utilisé dans les derniers cycles pour optimiser la qualité géométrique du modèle.

V.1.g. Logiciels de visualisation

Deux logiciels de visualisation ont été principalement utilisés au cours de cette thèse.

- **COOT (Emsley and Cowtan, 2004)**: Ce logiciel est utilisé pour la construction et la visualisation du modèle. Il est également utilisé pour calculer les différentes cartes de densité à partir des facteurs de structure contenus dans les fichiers MTZ. Il offre la possibilité de substituer les résidus, corriger et vérifier les angles Φ et Ψ , corriger les rotamères des chaînes latérales, construire et vérifier les molécules d'eau et ligands, corriger le modèle dans l'espace pour le faire correspondre à la densité. Il donne également de nombreuses possibilités pratiques (changement de nom des chaînes, numérotation de la chaîne, superposition de modèles). Le logiciel est régulièrement mis à jour et est notamment associé à Refmac depuis qu'il a intégré la suite CCP4 (version 6.0).
- **PyMoL (DeLano, 2002)** : Ce logiciel est principalement utilisé pour la visualisation des structures étudiées. Il propose différents types de représentation du modèle (cartoon pour visualiser graphiquement les structures secondaires, sphères où chaque atome est représenté par une sphère, ribbon pour observer uniquement le squelette polypeptidique). Il permet également de mettre en évidence des surfaces électrostatiques sur un modèle ou des sites actifs.

Ce logiciel est largement utilisé pour la création des différentes figures des structures étudiées présentées dans ce manuscrit (coloration des chaînes, ou d'une sélection définie, superposition). Contrairement à Coot, les étapes de construction et d'affinement sont ici impossibles.

V.2. Diffusion des rayons X aux petits angles (SAXS)

V.2.a. Principe de la diffusion des rayons X aux petits angles

La diffusion des rayons X aux petits angles est une technique permettant de déterminer des paramètres structuraux d'une molécule en solution (rayon de giration, D_{\max} , masse moléculaire, estimation de la forme globale, reconstruction de boucles flexibles à basse résolution). Comme la cristallographie, elle utilise la diffusion cohérente (Thomson) des rayons X. Les expériences de SAXS connaissent cependant certaines limites ; les macromolécules étant en agitation en solution, elles sont dans plusieurs orientations dans le faisceau et il peut également y avoir des conformations différentes en solution. De ce fait la diffusion observée est une moyenne sphérique (l'information 3D est perdue et réduite à une courbe de diffusion 1D) (**Figure 2.V.7**).

V.2.b. Enregistrement des données

Les données de SAXS utilisées au cours de cette thèse ont été obtenues sur la ligne de lumière BM29 de l'ESRF de Grenoble. En pratique, les données sont collectées sur un échantillon blanc (tampon) et sur la protéine. La courbe de diffusion analysée est le résultat de la soustraction de la diffusion du tampon à celle de la protéine. Pour les mesures, chaque échantillon (50 μ L) est soumis aux rayons X à quatre concentrations différentes et à 4°C en flux continu. Pour obtenir une moyenne et optimiser le rapport signal/bruit, 10 mesures sont effectuées par concentration (1 seconde d'exposition aux rayons X par mesure). Les concentrations utilisées pour chaque échantillon sont répertoriées dans le **tableau 2.V.4**.

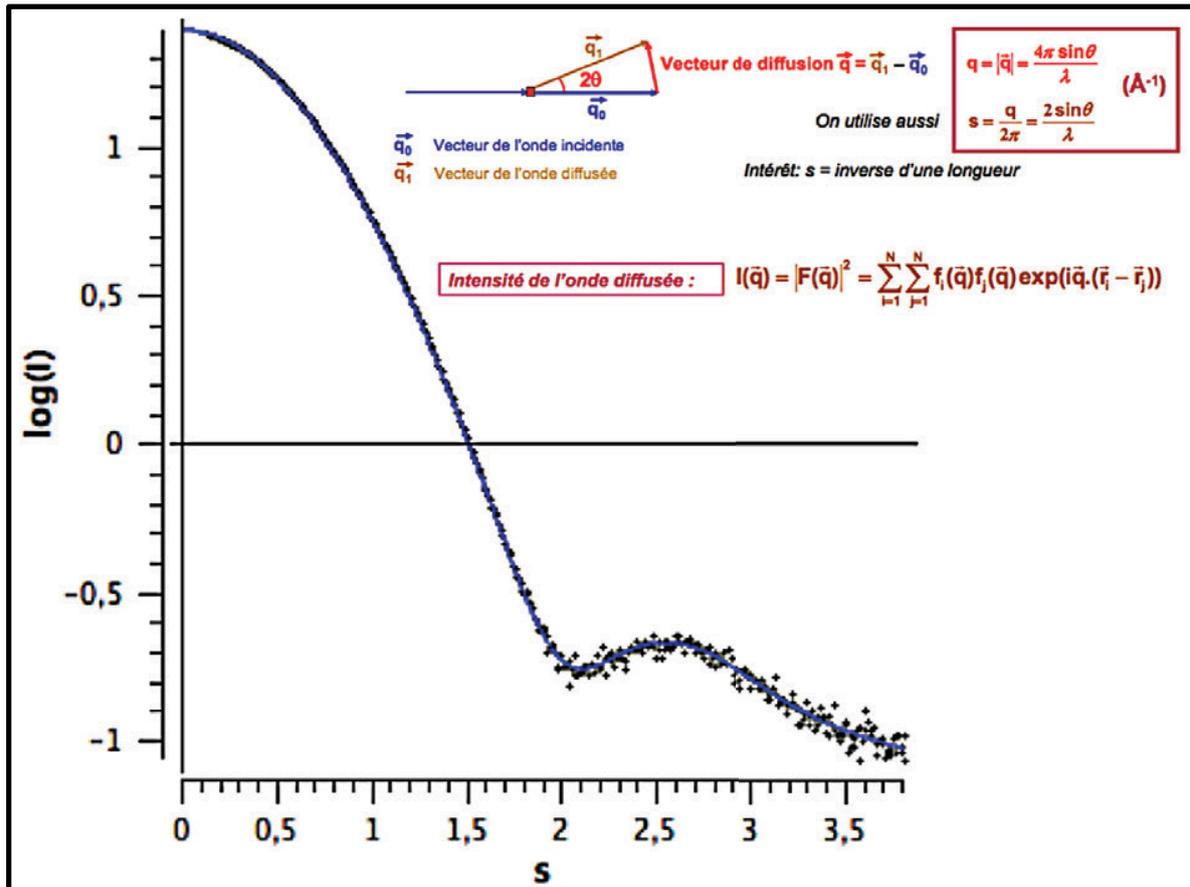


FIGURE 2.V.7: PRINCIPE DU SAXS

Le vecteur q est la soustraction du vecteur de l'onde diffusée au vecteur de l'onde incidente. L'intensité diffusée est enregistrée comme une fonction de momentum transfer q ou s ($4\pi \sin \theta / \lambda$ avec 2θ l'angle entre le rayon incident et la radiation diffusée). La distribution des intensités diffusées est isotropique à cause de l'orientation aléatoire des particules en solution (r). De ce fait l'intensité est maximale pour des valeurs de q ou s proches de 0 (domaine de Guinier). La diffusion pour un ensemble de protéines monodisperses n'interagissant pas entre elles en solution est proportionnelle à la diffusion d'une seule particule dans toutes les orientations.

En traçant le logarithme de l'intensité du rayon diffusé en fonction de q ou s , une courbe de diffusion est obtenue dont les régions nous renseignent sur les propriétés de l'objet en solution ($q=0$ détermination de la masse moléculaire, dans le domaine de Guinier la forme globale de la protéine, son rayon de giration, et dans le domaine des grands angles, la conformation de boucles, de domaines les uns par rapport aux autres).

| Molécules | Concentrations | Courbes enregistrées |
|----------------|----------------|----------------------|
| TcCRTFull | 1,63 mg/mL | 10/10 |
| | 3,09 mg/mL | 10/10 |
| | 7,05 mg/mL | 7/10 |
| | 10,4 mg/mL | 6/10 |
| TcCRTΔP | 1,17 mg/mL | 10/10 |
| | 4,48 mg/mL | 8/10 |
| | 8,35 mg/mL | 7/10 |
| | 12,15 mg/mL | 6/10 |
| TcCRTΔPΔ281 | 1,45 mg/mL | 10/10 |
| | 2,95 mg/mL | 10/10 |
| | 7,91 mg/mL | 10/10 |
| | 11,64 mg/mL | 8/10 |
| TcCRTΔPterΔ288 | 1,80 mg/mL | 10/10 |
| | 4,73 mg/mL | 9/10 |
| | 5,83 mg/mL | 10/10 |
| | 11,15 mg/mL | 6/10 |

TABLEAU 2.V.4: LISTE DES ECHANTILLONS ETUDIES PENDANT LES EXPERIENCES DE SAXS

Les 10 mesures à 3 concentrations permettent de valider l'expérience. En effet l'intensité $I(0)$ et le R_g ne doivent pas être dépendants du temps, ainsi il ne doit pas y avoir de différences entre les courbes au cours des 10 mesures pour chaque échantillon, sinon cela signifie la présence de dommages dus aux rayons X. Les dommages entraînent de l'agrégation et donc une augmentation de $I(0)$ et R_g aux petits angles. Les différentes concentrations permettent de juger des interactions. Pour être dans une situation monodisperse idéale, les valeurs $I(0)/C$ et R_g doivent être constantes, une augmentation avec la concentration indique une agrégation, une diminution avec la concentration indique des interactions répulsives entre les particules.

V.2.c. Traitement des données : Primus

La totalité des données SAXS obtenues au cours de cette thèse ont été traitées en utilisant les programmes contenus dans la suite **ATSAS (EMBL, Hambourg, (Petoukhov et al., 2012))**.

V.2.c.1. Mise à l'échelle

La première étape est la mise à l'échelle des courbes les unes par rapport aux autres en fonction des concentrations. Pour cela on se place dans la partie linéaire de la courbe (le domaine de Porod), où la diffusion n'est pas dépendante de la concentration. Si les concentrations ont été correctement mesurées le facteur de mise à l'échelle doit être de 1.

V.2.c.2. Fusion des données

Souvent les différentes régions de la courbe de diffusion ne sont pas de qualité égale aux différentes concentrations. En effet le rapport signal/bruit est souvent meilleur dans le domaine des grands angles pour une concentration importante. Au contraire le domaine de Guinier, grandement influencé par les interactions entre particules, est de très bonne qualité pour la concentration la plus basse. Avant d'interpréter les données il est donc intéressant de travailler avec une courbe optimale, en fusionnant les courbes de diffusion obtenues aux différentes concentrations. Pour cela, les régions d'intérêt sont définies pour chaque courbe et fusionnées avec l'outil « merge » du logiciel Primus.

V.2.c.3. Approximation de Guinier

Dans le cas d'une solution monodisperse idéale, il est possible de relier l'intensité diffusée à angle 0 à la masse moléculaire et au volume de la protéine en solution.

$$I(0) = V_{\text{sol}} \frac{cM}{N_a} \left(\frac{m_p N_a}{M} - \rho_s \bar{v}_p \right)^2$$

V_{sol} = volume de la solution
 M masse molaire de la particule, N_a nombre d'Avogadro
 m_p nombre d'e- d'une particule
 \bar{v}_p volume spécifique partiel de la particule

D'un point de vue pratique, la portion basse résolution de la courbe de SAXS n'est dictée que par un seul paramètre de taille. Ce paramètre, le rayon de giration (R_g), est la racine carrée des distances moyennes au carré de tous les diffuseurs par rapport au centre de la molécule.

A basse résolution, l'intensité diffusée peut-être décrite par l'approximation de Guinier, permettant d'exprimer $I(q)$ en fonction de R_g .

$$I(q) = I(0) \exp[-(q^2 R_G^2)/3]$$

Lorsque $qR_g < 1,3$ l'approximation de Guinier est valide pour ces valeurs de q . En pratique, $\ln(I)q = \ln I(0) - q^2 R_g^2 / 3$, donc en traçant le graphique $\ln(I)$ en fonction de q^2 nous obtenons une fonction linéaire dans le domaine des petits q , donc des grandes distances interatomiques à partir de laquelle R_g peut être déterminé et la masse moléculaire estimée à $I(0)$.

V.2.c.4. Kratky Plot

Le Kratky plot $I(q).q^2$ en fonction de q est une fonction pouvant être calculée directement depuis la courbe initiale et donnant une bonne indication sur le repliement de la protéine. La fonction d'une protéine bien repliée est une parabole suivie d'une asymptote vers 0, tandis que la position de cette parabole donne une bonne indication sur la taille de la protéine. Cette fonction permet d'évaluer la présence de domaines flexibles dans la protéine ; en effet, si de tels domaines sont présents, l'asymptote ne tendra pas vers 0 mais vers l'infini.

V.2.c.5. Fonction de distribution des distances

La courbe de diffusion d'un échantillon homogène est également directement dérivée de la distribution électronique de la molécule (c'est-à-dire de la fonction des distances interatomiques) $P(r)$. Pour interpréter une courbe de diffusion en terme de structure, il est nécessaire d'étudier la transformée de Fourier inverse de l'intensité pour obtenir la fonction de distribution des distances interatomiques $P(r)$. Cette fonction donne la fréquence à laquelle sont représentés les vecteurs de distances entre deux atomes.

$$P(r) = \frac{r^2}{2\pi^2} \int_0^{\infty} I(q) \frac{\sin(qr)}{qr} q^2 dq$$

Théoriquement, la fonction $P(r)=0$ lorsque la distance $r=0$ et si $r > D_{\max}$ (distance maximale entre les deux atomes les plus éloignés de la molécule, nous donnant ainsi le diamètre maximal de la molécule). La fonction $P(r)$ a un rôle important, car elle permet de calculer R_g et $I(0)$ en incluant toutes les données et n'est pas limitée à une petite région aux petits angles, comme l'approximation de Guinier. De ce fait, elle est moins sensible aux problèmes d'interaction ou d'agrégation. La fonction $P(r)$ est sensible à la symétrie et à la structure des domaines dans la protéine, et contient les informations de la forme de la protéine. Ainsi une protéine globulaire possèdera la majorité de ses atomes à égale distance du centre de la molécule, donnant une distribution gaussienne des distances.

La présence d'agrégats conduira à l'apparition de petits pics, pour des r élevés et le D_{\max} n'atteindra jamais 0, alors que des interférences entre protéines vont conduire à un D_{\max} trop court. Lors du traitement des données réelles, la fonction $P(r)$ est contrainte à 0 pour ces deux valeurs.

V.2.d. Génération d'enveloppe *in silico*

V.2.d.1. GASBOR

Le logiciel GASBOR est un logiciel capable de générer une enveloppe *ab initio*. La protéine est assimilée à un gaz d'atome, chaque atome est représenté par une perle qui sera contenue dans une sphère de diamètre D_{\max} . Puis la méthode du recuit simulé est utilisée afin d'obtenir un modèle atomique compact qui sera minimisé pour correspondre à notre courbe de diffusion. Cette méthode consiste à simuler une très haute température, qui fournit une énergie virtuelle au système et permet de lever les contraintes thermodynamiques, et donc permet au modèle de prendre des angles et torsions de liaison aberrants. L'énergie de la molécule est calculée, puis un changement de conformation est effectué, et la différence entre la nouvelle énergie et l'ancienne est recalculée : si elle est inférieure à 0, la nouvelle configuration est acceptée. Le calcul est répété en diminuant progressivement la température virtuelle pour réintroduire les contraintes thermodynamiques, jusqu'à atteindre un modèle d'enveloppe correspondant à un minimum énergétique, et donc pouvant correspondre à un bon modèle, car stable.

Le modèle généré doit être en adéquation avec la courbe de diffusion et donc la fonction $P(r)$ calculée. Il est nécessaire de générer plusieurs enveloppes indépendantes, afin de les comparer entre elles pour vérifier la robustesse du modèle.

V.2.d.2. DAMAVER

Le logiciel DAMAVER permet de comparer et combiner les enveloppes générées par GASBOR pour aboutir à un modèle final en 3 étapes (composé de 3 sous-logiciels) :

Damsel: vérifie que les fichiers PDB sont bien présents, les compare entre eux et génère une matrice (si on vérifie le fichier `damsel.log` et que la déviation (NSD) est proche de 1, les enveloppes sont proches).

Damsup: utilise le fichier de sortie de Damsel, et décide si un fichier est trop différent et doit être exclu.

Damaver: combine tous les fichiers PDB préalablement inclus, pour en faire une enveloppe moyenne.

Un certain recul est nécessaire lors de l'analyse des enveloppes. En effet le SAXS ne fournit qu'une information unidimensionnelle (1D), et les informations 3D présentées ne sont que déduites par le calcul de la courbe de diffusion.

V.2.e. Structure cristallographique vs Structure en solution

V.2.e.1. Crysol

Le programme Crysol génère une courbe de diffusion théorique à partir d'un modèle PDB (structure obtenue par RMN, cristallographie aux rayons X, ou modélisation..) et la compare avec la courbe de diffusion expérimentale mesurée lors de l'expérience SAXS. La valeur de CHI^2 doit être la plus proche de 1 et, d'un point de vue graphique, les deux courbes doivent être superposables.

V.2.e.2. BUNCH

Le programme BUNCH combine une approche de minimisation par corps rigide et la modélisation *ab initio* pour modéliser des protéines multi-domaines ou oligomériques. Il se base sur les courbes de diffusion pour déplacer dans l'espace des unités rigides (par exemple des modèles issus de fichier PDB) et modéliser des parties manquantes (comme des boucles) reliant ces corps rigides. Il utilise ensuite le programme Crysol pour comparer le modèle créé avec les données expérimentales. Par exemple, dans l'étude d'un complexe, le logiciel Bunch utilisera les courbes de diffusion des deux protéines individuelles constituant le complexe et celle du complexe, ainsi que le modèle PDB de chacune de deux protéines pour réarranger les deux modèles de manière à correspondre au mieux à la courbe de diffusion du complexe.

V.3. Modélisation de la dynamique de la protéine

Les différentes charnières définissant des régions flexibles dans les calréticulines ont été déterminées en utilisant le serveur Elnemo (**Suhre and Sanejouand, 2004**). Ce serveur utilise l'analyse des modes normaux pour prédire les mouvements possibles d'une macromolécule en solution.

V.4. Résonance Magnétique Nucléaire

V.4.a. Principe

La RMN repose sur les propriétés des particules sub-atomiques (proton, neutron, électron) et notamment le spin de ces particules. Selon la règle d'exclusion de Pauli, dans une orbitale on ne peut placer que 2 électrons qui ont un spin forcément opposé. Le spin est un moment cinétique intrinsèque, c'est une grandeur vectorielle. Les noyaux composés d'un nombre pair de protons et de neutrons ont un nombre quantique de spin nul et ne possèdent donc pas de spin nucléaire. Pour avoir un signal à mesurer, le spin de l'atome doit être de 1/2, c'est à dire que un neutron ou un proton ne doit pas être apparié (c'est le cas de ^1H , ^{13}C ou ^{15}N). Le spin est également associé et proportionnel à un vecteur magnétique colinéaire.

En appliquant un champ magnétique à ces atomes de spin et de moment magnétique non nul, le noyau absorbe cette énergie, puis va revenir à l'équilibre (relaxation magnétique longitudinale) selon une certaine fréquence qui est mesurée. Plus simplement, lors de l'application du champ magnétique, les moments magnétiques de chaque atome qui sont normalement orientés dans toutes les directions vont essayer de s'aligner avec le champ magnétique selon un mouvement de précession (précession de Larmor) possédant une certaine fréquence qui sera mesurée.

Cette fréquence de retour à l'équilibre est influencée par l'environnement du noyau étudié (notion de blindage). C'est ce principe que nous avons utilisé lors de l'étude de la formation du complexe, en observant l'impact du changement d'environnement par l'ajout d'un second partenaire sur les noyaux sondés.

Lors d'une expérience RMN, la fréquence de retour à l'équilibre est exprimée sous forme de déplacement chimique. Ce déplacement chimique correspond au rapport entre la fréquence de retour à l'équilibre du noyau étudié à laquelle est soustraite la fréquence d'un noyau de référence et la fréquence de travail issue de la combinaison du champ magnétique généré par l'aimant et du rapport gyromagnétique du noyau étudié).

V.4.b. Enregistrement des spectres 1D ^1H et 2D ^1H ^{15}N

Les spectres RMN 1D du proton (corrélation entre protons voisins) et 2D ^1H ^{15}N (corrélation entre proton et azote liés) ont été mesurés et interprétés sur la plateforme RMN de l'IBS par Adrien Favier. Les spectres RMN ont été enregistrés à 25°C sur le spectromètre BRUKER

700 MHz, équipé d'une sonde cryogénique inverse triple résonance 5 mm TCI 1H/15N/13C. Les analyses RMN ont été enregistrées sur des échantillons dissous dans 600 μ L de H₂O-D₂O à 95-5%, en utilisant les expériences à 2 dimensions SOFAST et BEST-TROSY développées à l'IBS ((Favier and Brutscher, 2011; Schanda and Brutscher, 2005)). Les spectres ont été obtenus avec des délais de 200 ms entre chaque scan. Les déplacements chimiques sont exprimés en ppm par rapport au DSS comme référence externe.

VI. Analyse biochimique des interactions entre les protéines étudiées

VI.1. ELISA

VI.1.a. Principe

Les tests ELISA (Enzyme-linked ImmunoSorbent Assay), souvent utilisés en immunologie, permettent d'étudier (qualitativement ou quantitativement) les interactions entre deux protéines, ou de détecter la présence d'un antigène ou d'un anticorps dans un sérum.

Une première protéine (un partenaire d'une interaction, ou un anticorps dirigé contre une protéine à identifier dans un sérum) est adsorbée sur une surface plastique (couramment au fond d'un puits d'une plaque de 96 puits) et les sites libres sont saturés par une protéine neutre (ne devant pas interagir dans la réaction, comme la BSA). Le second partenaire (ou un sérum contenant l'antigène à identifier) est ajouté au contact de la protéine adsorbée et, après plusieurs lavages qui visent à dissocier les interactions faibles ou non spécifiques, la présence du second partenaire est révélée par un anticorps spécifique. Cette révélation peut être colorimétrique, fluorescente ou radioactive et peut être directe (l'anticorps dirigé contre la protéine est directement couplé à une molécule permettant la révélation) ou indirecte (comme dans le cas du Western Blot où un second anticorps dirigé contre l'isotype du premier est lui-même couplé à une molécule colorée, une enzyme ou un fluorophore).

En utilisant des concentrations de protéines croissantes, cette méthode permet également de déterminer des constantes d'affinité pour une interaction donnée, même si, contrairement à la technique de SPR, l'affinité calculée sera ici le reflet de l'affinité de l'interaction entre les partenaires, mais aussi entre l'anticorps primaire et le second partenaire, et entre l'anticorps secondaire et l'anticorps primaire. La valeur obtenue peut aussi être influencée par l'efficacité de la réaction qui dépend de nombreux paramètres (utilisateur, température, etc.). Il est

souvent nécessaire d'optimiser également les temps et températures d'incubation et de révélation, la nature des tampons ou des protéines neutres utilisées qui ne doivent avoir d'influence ni sur l'interaction, ni sur la réaction de révélation.

VI.1.b. Protocole de mesure de l'interaction entre les fragments de calréticulines et les régions de C1q

En nous basant sur des expériences existantes dans la littérature pour caractériser les interactions entre les différentes régions de C1q et les calréticulines de *T. cruzi* et *E. histolytica* par ELISA, nous avons établi un protocole nous permettant de mettre en évidence les interactions entre les différents fragments de calréticulines et les régions de C1q par ELISA (Ferreira et al., 2004a; Ramírez et al., 2011a; Vaithilingam et al., 2012).

Après chacune des étapes citées, les puits sont lavés quatre fois avec 200 µL d'un **tampon de lavage PBS-T (NaCl 137 mM, KCl 2,7 mM, Na₂HPO₄ 10 mM, KH₂PO₄ 1,76 mM pH 7,4 Euromedex + 0,05% Tween 20)**

- Les différents fragments de calréticulines sont dilués aux concentrations souhaitées (**Tableau 2.VI.1**) dans du tampon PBS et immobilisés sur une plaque de 96 puits Maxisorb Nunc-Immuno F96 (Nunc) pendant une nuit à 4°C sous agitation.
- Les puits sont saturés avec 200 µL du **tampon de blocage (PBS+BSA 2%)** et la plaque est incubée 2 h à 37°C en chambre humide.
- C1q (10 nM), les régions globulaires de C1q (GRC1q) (10 nM) et les régions collagènes de C1q (CLR) (200 nM) sont dilués dans du tampon **PBS+BSA 0,2%+Tween 20 0,05%**, 100 µL sont ajoutés dans chaque puits, et la plaque est incubée 2 h à 37°C en chambre humide.
- La fixation de C1q, des CLR et des GRs est détectée par ajouts successifs de 100 µL d'un anticorps polyclonal anti-C1q (produit chez le lapin, dilution 1/1000) ou anti-GR (produit chez le lapin, dilution 1/500), suivi d'un anticorps anti-IgG de lapin couplés à la HRP (produit chez la chèvre, dilution 1/10000). Les anticorps sont dilués dans du tampon **PBS+BSA 0,2%+Tween 20 0,05%**. A chaque étape, la plaque est incubée 1 h à 37°C en chambre humide.
- La révélation (coloration bleue) s'effectue par ajout du substrat spécifique de la HRP (3,3',5,5'-tetramethylbenzidine (TMB) (Sigma Aldrich) (200 µL)) pendant 20 minutes

à 20°C. La réaction est stoppée par ajout de 100 µL de H₂SO₄ 0,5 M et l'absorbance des échantillons, qui est proportionnelle à la quantité de C1q fixée, est lue à 450 nm avec un lecteur de plaques (Victor™ Multilabel Counter 1420, Wallac).

Les valeurs indiquées pour chaque point correspondent à la moyenne de 6 mesures (2 expériences indépendantes réalisées en triplicate) et chaque valeur initiale a été corrigée par la soustraction des interactions non spécifiques des protéines dans les puits ne contenant aucune protéine adsorbée.

| Echantillons (ligands) | Concentrations Max (facteur de dilution) |
|------------------------|--|
| TcCRTFull (C1q-GR-CLR) | 1,2 µM(1/2)-0,6 µM(1/3)-1 µM(1/2) |
| TcCRTΔP (C1q-GR-CLR) | 0,5 µM(1/2)-0,3 µM(1/3)-0,8 µM(1/2) |
| TcCRTCrys (C1q-GR-CLR) | 0,5 µM(1/2)-0,25 µM(1/2)-0,5 µM(1/2) |
| TcCRTNGB (C1q-GR-CLR) | 0,2 µM(1/2)-0,2 µM(1/2)- 0,2 µM(1/2) |
| EhCRTΔ288 (C1q-GR-CLR) | 0,18 µM(1/2)-0,18 µM(1/2)- 0,18 µM(1/2) |
| EhCRTCrys (C1q-GR-CLR) | 0,06 µM(1/2)-0,06 µM(1/2)- 0,12 µM(1/2) |
| EhCRTNGB (C1q-GR-CLR) | 0,12 µM(1/2)-0,05 µM(1/2)- 0,05 µM(1/2) |
| HuCRTFull (C1q-GR-CLR) | 0,3 µM(1/2)-0,3 µM(1/2)-0,6 µM(1/2) |
| HuCRTNGB (C1q-GR-CLR) | 0,6 µM(1/2)-0,3 µM(1/2)-0,2 µM(1/2) |

TABLEAU 2.VI.1: CONCENTRATIONS DE CRTS ADSORBÉES POUR LES TESTS D'INTERACTIONS ELISA

Dans ces expériences, chaque fragment est incubé à 8 concentrations différentes avec les ligands correspondants (C1q, GR, CLR). Les concentrations sont obtenues par dilution en cascade à partir de la concentration maximale, en utilisant le facteur de dilution indiqué entre parenthèse.

VI.1.c. Mise en place d'un protocole testant l'inhibition de la voie classique du complément par les différents fragments de calréticulines

En nous basant sur les expériences menées par nos collaborateurs à Santiago (**Ferreira et al., 2004a; Valck et al., 2010**), pour déterminer l'impact de la calréticuline de *T. cruzi* sur l'activation de la voie classique du complément par ELISA, un protocole a été mis au point pour vérifier l'effet des différents fragments de calréticuline de *T. cruzi* sur cette inhibition. Après de nombreuses mises au point, les résultats observés dans la littérature n'ont pas été

reproduits mais le protocole exposé a permis l'obtention de contrôles solides et servira de base pour vérifier l'effet de ces différents fragments.

Comme décrit dans l'introduction, l'activation de la cascade de la voie classique du complément débute initialement par la protéolyse par le complexe C1 de la protéine C4 en fragments C4a et C4b. La protéine C4b a la capacité de se fixer de manière covalente au fond des puits des plaques de 96 puits. La quantité de molécule C4b déposée au fond du puits est directement proportionnelle à l'activité de la voie classique du complément et pourra être mise en évidence par un test ELISA classique décrit dans la section précédente.

Avant le début de l'expérience, toutes les protéines (fragments de calréticulines, protéines neutres) sont dialysées contre le tampon **VB (5,5 acide diéthylbarbiturique 5 mM, NaCl 140 mM, pH 7,4)**.

Après chacune des étapes citées, les puits sont lavés quatre fois avec 200 µL d'un **tampon de lavage PBS-T**.

- 0,2 µg d'IgM (Sigma Aldrich) sont adsorbés sur une plaque de 96 puits Maxisorb Nunc-Immuno F96 (Nunc) sur la nuit à 4°C dans du tampon carbonate (**Na₂CO₃ 15 mM NaHCO₃ 35 mM pH 9,6**). Un puits témoin sans IgM (uniquement du tampon) est réalisé.
- Les puits sont saturés avec 200 µL de **tampon de blocage (PBS+BSA 3%)** et la plaque est incubée 1 h à 37°C en chambre humide.
- Du sérum humain normal (SHN) est dilué au 1/450^e (0,3%) dans un volume souhaité pour l'expérience de tampon **VB contenant 5 mM de CaCl₂ et 1,5 mM de MgCl₂ pH 7,4 (VB⁺⁺)**, en présence de 10 µM des différents fragments de calréticuline, de BSA ou d'ovalbumine. Les échantillons sont incubés à 37°C pendant 30 min. Un témoin sans sérum (uniquement du tampon VB⁺⁺) est réalisé.
- Après incubation, les échantillons sont mis au contact des IgM (100 µL/puits) et la plaque est incubée 1 h à 37°C en chambre humide.
- La fixation de C4b est détectée par ajout de 100 µL d'un anticorps polyclonal anti-C4 (produit chez le lapin dilution 1/1000, Siemens), suivi d'un anticorps anti-IgG de lapin, couplé à la HRP (produit chez la chèvre, dilution 1/20000). Les anticorps sont dilués dans du tampon **PBS+BSA 0,2%+Tween 20 1%**. A chaque étape, la plaque est incubée 1 h à 37°C en chambre humide.

- La révélation (coloration bleue) s'effectue par ajout du substrat spécifique de la HRP (TMB (Sigma Aldrich) (200 µL)) pendant 6 minutes à 20°C. La réaction est stoppée par ajout de 100 µL de H₂SO₄ 0,5 M et l'absorbance des échantillons, qui est proportionnelle à la quantité de C1q fixée, est lue à 450 nm avec un lecteur de plaques (VictorTM Multilabel Counter 1420, Wallac).

Les valeurs indiquées pour chaque point correspondent à la moyenne de 6 mesures (2 expériences indépendantes réalisées en triplicate) et chaque valeur initiale a été corrigée par la soustraction des interactions non spécifiques des anticorps primaires dans les puits ne contenant ni IgM adsorbée, ni sérum (donc normalement sans fixation de molécule C4b).

VI.2. Résonance plasmonique de surface (SPR)

La SPR est une technique basée sur un phénomène physique utilisée pour mesurer la liaison d'une molécule en solution sur une molécule immobilisée à la surface d'une couche métallique. Elle permet une étude en temps réel sans marquage.

Dans les métaux comme l'or, il existe des ondes longitudinales particulières, les plasmons, correspondant à une oscillation des densités de charges à l'interface de deux milieux possédant des indices de réfraction différents (verre et liquide par exemple). La traversée du premier milieu (verre) par un rayon monochromatique (laser) à une longueur d'onde choisie pour obtenir sa réflexion totale (réflexion totale interne), crée une onde évanescente concomitante qui se propage dans le milieu de plus faible indice de réfraction. Sous des angles incidents particuliers, les plasmons peuvent être excités par le rayon monochromatique (interaction avec les électrons libres de l'or par exemple) et entrer en résonance avec l'onde évanescente. L'absorption de cette énergie conduit à une diminution de l'intensité du rayon réfléchi. L'angle pour lequel l'intensité est la plus faible s'appelle « angle de résonance plasmonique ». L'énergie nécessaire à l'excitation des plasmons est directement liée à l'indice de réfraction du second milieu, qui peut être modifié par exemple avec la fixation de molécules biologiques à l'interface.

La technologie Biacore[®], repose sur le principe de la SPR. Une première molécule, le ligand, est immobilisé sur une fine couche d'or déposée sur une surface de verre appuyée contre un prisme semi-cylindrique. Un rayon monochromatique est focalisé sur la surface en verre et

des capteurs mesurent l'angle et l'intensité des faisceaux réfléchis en prenant comme référence l'angle de résonance plasmonique. La fixation d'une seconde molécule (analyte) injectée dans un flux de tampon sur le ligand entraîne une modification de l'indice de réfraction du milieu à proximité des deux surfaces, entraînant une modification de l'angle. Cette modification est proportionnelle à la masse fixée sur la surface et se traduira en un signal exprimé en RU (résonance unit), **Figure 2.VI.1**. Un signal de 1 RU correspond à une variation d'angle de 10^{-4} degré, correspondant à une variation d'indice de réfraction de 10^{-6} et à une masse de 1pg fixée par mm^2 de surface (sur une puce de type CM5).

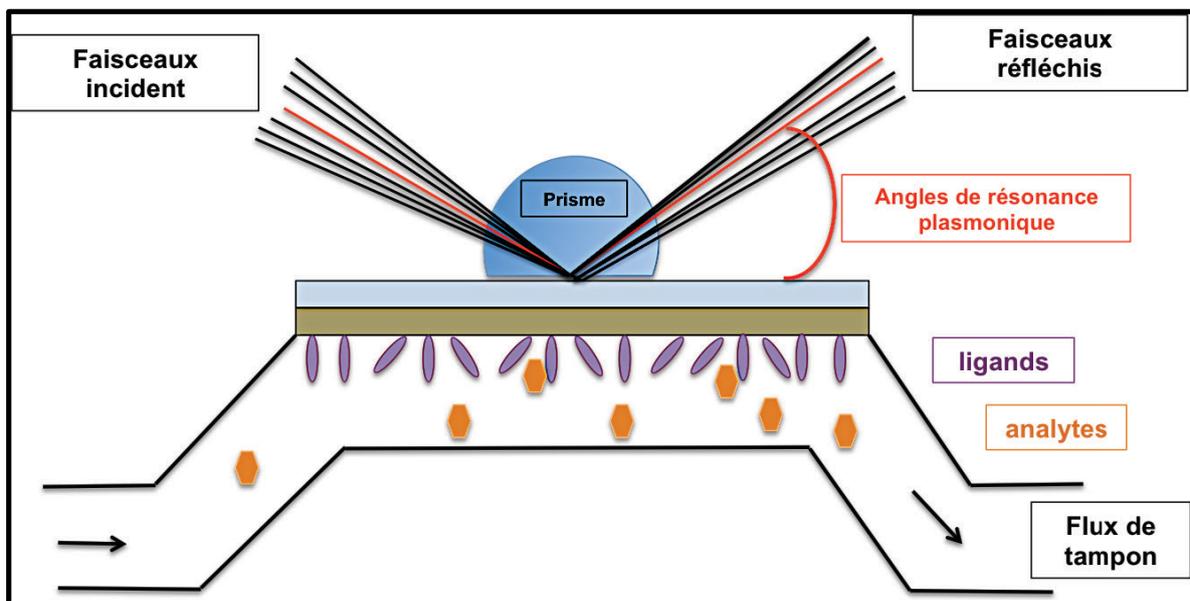


FIGURE 2.VI.1 : PRINCIPE DE LA RESONANCE PLASMONIQUE DE SURFACE (ADAPTE DU MANUEL BIACORE SENSOR SURFACE HANDBOOK)

Le sensorgramme permet de suivre l'interaction en temps réel. L'analyte est injecté sur la surface exposant le ligand immobilisé et s'associe progressivement (phase d'association) jusqu'à l'équilibre (Req). L'injection d'analytes en continu augmente progressivement la réponse jusqu'à atteindre R_{max} , correspondant à la capacité maximale de fixation de la surface. Après arrêt de l'injection, le flux continu de tampon permet la dissociation des analytes (phase de dissociation). La vitesse de dissociation est dépendante de l'affinité de l'interaction. Pour les interactions les plus fortes, une étape de régénération est nécessaire pour revenir au signal de base (essentielle avant une nouvelle injection). Les phases d'association et de dissociation permettent de déterminer les constantes de vitesses d'association (k_a) et de dissociation (k_d), utilisées pour calculer la constante apparente de dissociation à l'équilibre (K_D), (**Figure 2.VI.2**).

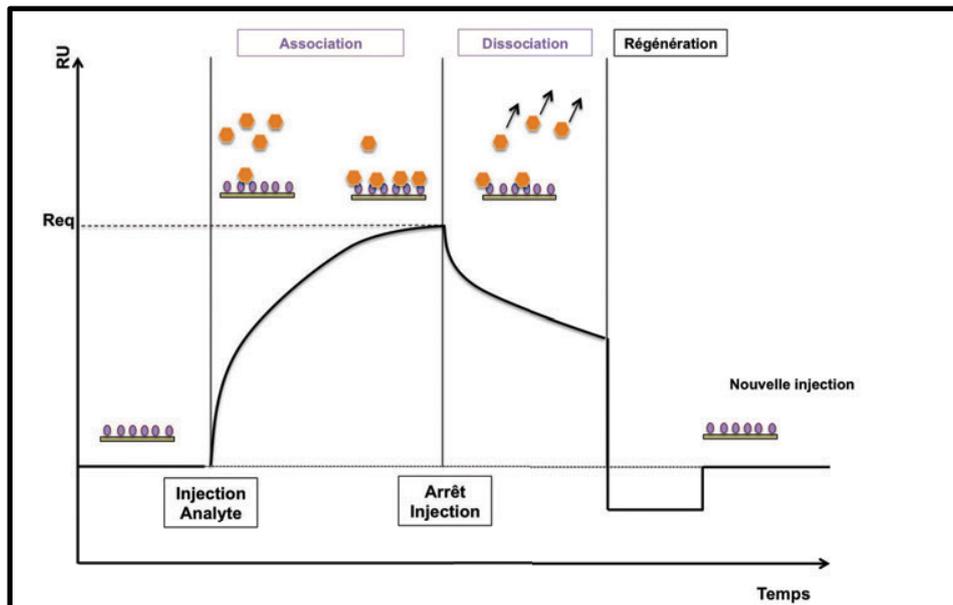


FIGURE 2.VI.2 : DIFFERENTES PHASES D'UN SENSORGRAMME (ADAPTE DU MANUEL BIACORE, SENSOR SURFACE HANDBOOK)

VI.2.a. Immobilisation des ligands sur la surface

Les expériences de SPR sont effectuées sur les Biacore T200 et 3000 (GE Healthcare Life Sciences Technologies) à 25°C sur des puces (sensorchips) présentant 4 canaux distincts (flow-cell Fc), sur lesquels peuvent être immobilisés des ligands différents. Le système de micro fluidique permet de sélectionner les pistes sur lesquelles nous souhaitons travailler.

Pour interagir de manière covalente avec la surface chargée négativement, les protéines doivent s'en rapprocher et donc être chargées positivement. Pour cela, les échantillons sont dilués dans un tampon dont le pH est inférieur au point isoélectrique de la protéine (pI). Pour déterminer le meilleur tampon auquel travailler, un « pH scouting » est effectué, ce qui consiste à injecter les différentes protéines diluées dans des tampons aux pH souhaités sans activation de la surface (pour que les interactions soient uniquement électrostatiques) et observer les phases d'association et de dissociation du signal. Le pH choisi est celui présentant une fixation importante à une vitesse moyenne (une vitesse d'association trop importante entraîne des injections de faible volume par le programme d'immobilisation, qui s'ajustent moins bien au niveau d'immobilisation souhaité) et une dissociation permettant de revenir à la ligne de base.

La surface d'or, décorée de groupements dextrans carboxyméthylés (« sensor Chip CM5 » (Research Grade, GE Healthcare)), est activée avec un mélange 1 : 1 (80 µL) de 1-ethyl-3-(3-

dimethylpropyl)-carbodiimide 0,4 M (EDC) et de N-Hydroxysuccinide 0,1 M (NHS) formant des esters succinimide de haute réactivité capables de former des liaisons covalentes avec les amines primaires des protéines.

Après dilution des protéines dans les tampons sélectionnés, l'injection s'effectue dans un flux de tampon **HBS-P (Hepes 10 mM, NaCl 145 mM, surfactant P20 0,05% pH 7,4)** à un débit de 10 $\mu\text{L}/\text{min}$ pour atteindre une réponse (RU) prédéfinie. Pour comparer l'interaction d'un analyte sur différents ligands, il est nécessaire que le nombre de molécules fixées sur la surface pour chaque ligand soit équivalent. En se basant sur les travaux d'interactions effectuées sur la calréticuline humaine (**Païdassi et al., 2011**), la réponse souhaitée pour chaque protéine immobilisée est choisie en fonction de sa masse moléculaire (**Tableau 2.VI.2**). A la fin de l'injection, les esters libres sont bloqués par l'injection de 80 μL d'éthanolamine 1 M (pH 8,5).

Une surface libre (généralement le canal Fc1) est utilisée comme surface de référence (aucun ligand n'est immobilisé). Elle est activée par le mélange EDC/NHS et tous les esters formés sont directement bloqués par l'éthanolamine. Son signal représente les interactions non spécifiques et sera soustrait du signal des autres pistes.

| Protéines immobilisées | Tampon de dilution (10 mM acétate de sodium) | Concentrations injectées | Masses Moléculaires (Da) | Quantité moyenne visée (RU) | Quantité immobilisée (RU) |
|-------------------------------|---|---------------------------------|---------------------------------|------------------------------------|----------------------------------|
| HuCRTFull | pH 4 | 20 $\mu\text{g}/\text{mL}$ | 49136,5 | 2456 | 3330 |
| EhCRTA288 | pH 4 | 20 $\mu\text{g}/\text{mL}$ | 41025,6 | 2906 | 2800 |
| TcCRTFull | pH 4 | 20 $\mu\text{g}/\text{mL}$ | 44255,1 | 3000 | 3080 |
| TcCRTCrys | pH 4,5 | 20 $\mu\text{g}/\text{mL}$ | 31111,9 | 2100 | 2037 |
| EhCRTCrys | pH 4 | 20 $\mu\text{g}/\text{mL}$ | 31425,1 | 2121 | 2210 |
| HuCRTNGB | pH 4,5 | 20 $\mu\text{g}/\text{mL}$ | 33328,1 | 2250 | 2357 |
| EhCRTNGB | pH 4,5 | 20 $\mu\text{g}/\text{mL}$ | 30442,1 | 2054 | 2060 |
| TcCRTNGB | pH 4,5 | 20 $\mu\text{g}/\text{mL}$ | 30118,8 | 2032 | 2097 |

TABLEAU 2.VI.2: CONDITIONS D'IMMOBILISATION DES PROTEINES ETUDIEES (LA QUANTITE IMMOBILISEE DE HUCRT EST UTILISEE COMME REFERENCE)

VI.2.b. Injection des analytes

Les protéines dont on souhaite analyser l'interaction avec les ligands immobilisés sont diluées dans du tampon de course **HBS-P** et injectées dans ce même tampon à un débit de 30 $\mu\text{L}/\text{min}$ pendant 140 s avec une période de dissociation de 180 s. Si tous les complexes ne sont pas dissociés après la phase de dissociation, la surface est régénérée par des ajouts de 10 à 20 μL de NaOH (10 à 20 mM).

VI.2.c. Détermination des paramètres cinétiques des interactions

Pour déterminer les paramètres cinétiques, des concentrations croissantes d'analytes sont injectées sur un même ligand immobilisé grâce à un programme établi, sans étape de régénération entre chaque injection (Single Cycle Kinetics SCK). Les différentes courbes sont corrigées par soustraction du signal de la piste de référence, ainsi que par le signal correspondant à l'injection du tampon seul. Les différents paramètres cinétiques (k_a , K_D , k_d) et le Rmax (capacité de fixation maximale) sont calculés à partir de cinq concentrations avec le logiciel Biacore T200 Evaluation Software version 1, en utilisant le modèle Langmuir 1 : 1 ou un modèle 1 : 1 impliquant un changement de conformation (dans ce modèle, le complexe initialement formé change de conformation pour évoluer vers un complexe plus stable). Le paramètre χ^2 mesure la qualité de l'ajustement des sensorgrammes expérimentaux aux sensorgrammes théoriques obtenus pour le modèle choisi.

3. RESULTATS PRINCIPAUX

Chapitre 1 : Analyse structurale des calréticulines de *T. cruzi* et *E. histolytica*

La récente résolution de la structure du domaine globulaire de la calréticuline humaine et de la structure incluant le domaine globulaire et un fragment du domaine P de la calréticuline de souris en complexe avec un tétrasaccharide (Chouquet et al., 2011) ont permis de mettre en évidence une importante conservation du domaine lectine contenant le site de fixation spécifique du carbohydrate, ainsi qu'un nouveau site de fixation potentiel des peptides pouvant participer au repliement des protéines non-glycosylées (Kozlov et al., 2010a).

Mon projet de thèse a débuté au centre d'une collaboration établie entre notre laboratoire et l'équipe du Dr. Ferreira à l'université de médecine de Santiago au Chili. Son équipe s'intéressait à la calréticuline du parasite *T. cruzi* possédant des propriétés d'interactions avec les collagènes de défense dont C1q. Nous nous sommes donc proposé de mettre à profit notre savoir faire dans le domaine de la biologie structurale, pour résoudre la structure à haute résolution de la calréticuline de *T. cruzi*.

Grâce au travail fourni successivement par Mickaël Jacquet, Emmanuelle Laffly, Anne Chouquet et Marina Ianello avant mon arrivée au laboratoire, j'ai pu démarrer ce projet dans un contexte où les protocoles de production et de purification de la calréticuline de *T. cruzi* venaient d'être mis au point et maîtrisés.

I. Purification des fragments de calréticulines

I.1. Stratégie de dissection

Dans le but d'obtenir un fragment cristallisable, nous avons choisi de tronquer les domaines flexibles et chargés (domaine P, hélice C-terminale (Figure 3.I.1 (A))) pouvant gêner l'empilement cristallin. Lors de mon arrivée au laboratoire, des essais de cristallogénèse avaient été entrepris sans succès sur la protéine entière et sur un fragment possédant un domaine P raccourci d'après la structure de la calréticuline de souris, comme illustré sur la figure 3.I.1 (Pocanschi et al., 2011).

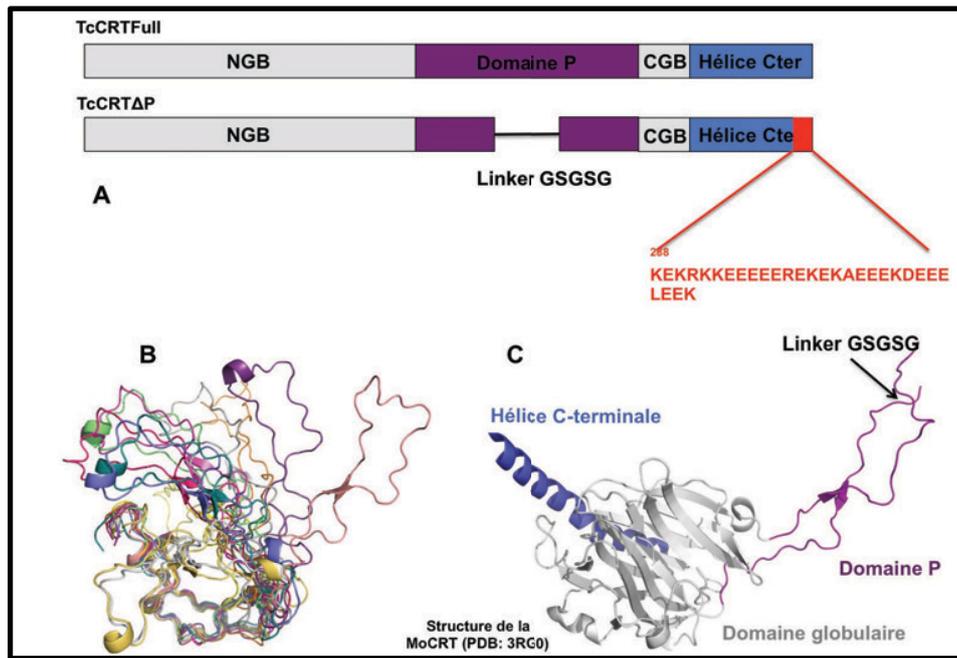


FIGURE 3.I.1 : DESCRIPTION DE LA CONSTRUCTION DE TcCRT Δ P

(A) Pour obtenir un fragment de calréticuline capable de cristalliser, il est nécessaire de réduire l'importance des régions flexibles constituées par le domaine P (magenta) et l'hélice C-terminale (bleu). La séquence très chargée de l'extrémité C-terminale est présentée en rouge. (B) L'étude RMN du domaine P confirme son importante flexibilité (Elgaard et al., 2001a). (C) La structure de la MoCRT représente le fragment de calréticuline cristallisé le plus long (PDB : 3RG0) (Pocanschi et al., 2011). La position choisie pour introduire le linker dans notre construction Δ P est indiquée par la flèche noire.

Ces échecs répétés nous ont donc conduit à nous intéresser à une seconde région flexible et chargée : l'hélice C-terminale. Dans un premier temps, nous avons testé plusieurs coupures de l'hélice C-terminale comme cela avait déjà été réalisé avec succès dans le cas de l'étude de la HuCRT, puis des délétions supplémentaires ont été successivement réalisées au voisinage de résidus potentiellement impliqués dans la stabilisation de l'hélice contre le domaine globulaire (Figure 3.I.2).

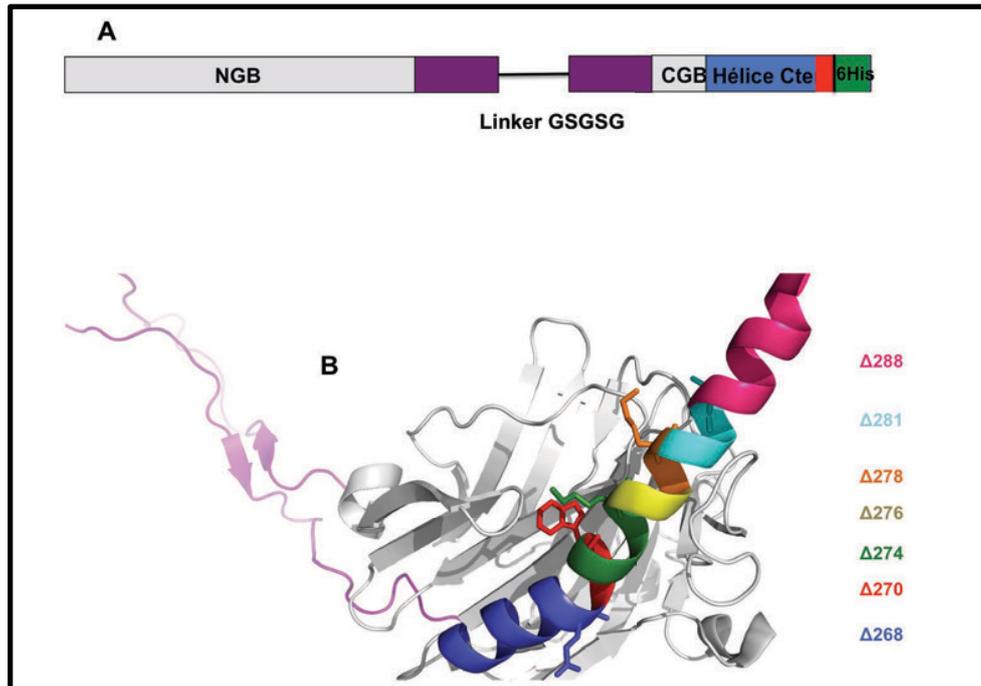


FIGURE 3.I.2 : STRATEGIE DE DISSECTION DE L'HELICE C-TERMINALE

(A) L'extrémité de l'hélice C-terminale (**en rouge**), composée de nombreux résidus chargés, a initialement été supprimée (après le résidu 288). (B) L'hélice a ensuite été successivement tronquée selon le schéma décrit. Le nombre associé à chaque construction correspond à la numérotation du dernier acide aminé précédant la coupure.

Les acides aminés correspondant aux régions de la protéine à supprimer ont été enlevés en utilisant un protocole de mutagenèse dirigée, décrit plus précisément dans la section Matériels et Méthodes. Comme décrit, les amorces ont été choisies pour s'hybrider sur la séquence codant pour la TcCRT en amont de la délétion, et sur la séquence codant pour l'étiquette 6-Histidines située directement en aval de chacune des coupures. Les plasmides contenant les gènes codant pour les différentes constructions des fragments de la TcCRT ont été vérifiés par séquençage (alignements non présentés) et utilisés pour transformer les souches *E. coli BL21 (DE3)* dans le but de produire en grande quantité les protéines d'intérêt.

I.2. Production des fragments de TcCRT

Les différents fragments de la TcCRT (TcCRT Δ P Δ 268, 270, 274, 276, 278, 281 et 288) ont été produits en suivant le protocole décrit dans la section Matériels et Méthodes. L'analyse sur gel SDS-PAGE des fractions de protéines totales avant et après induction à l'IPTG révèle une bande de forte intensité migrant à une distance correspondant à la masse apparente de chacun des fragments (33,2 kDa, 32,7 kDa, 32,5 kDa, 32,3 kDa, 32 kDa et 31 kDa, respectivement). Les gels présentés sur la **figure 3.I.3** mettent en évidence leur surproduction conséquente.

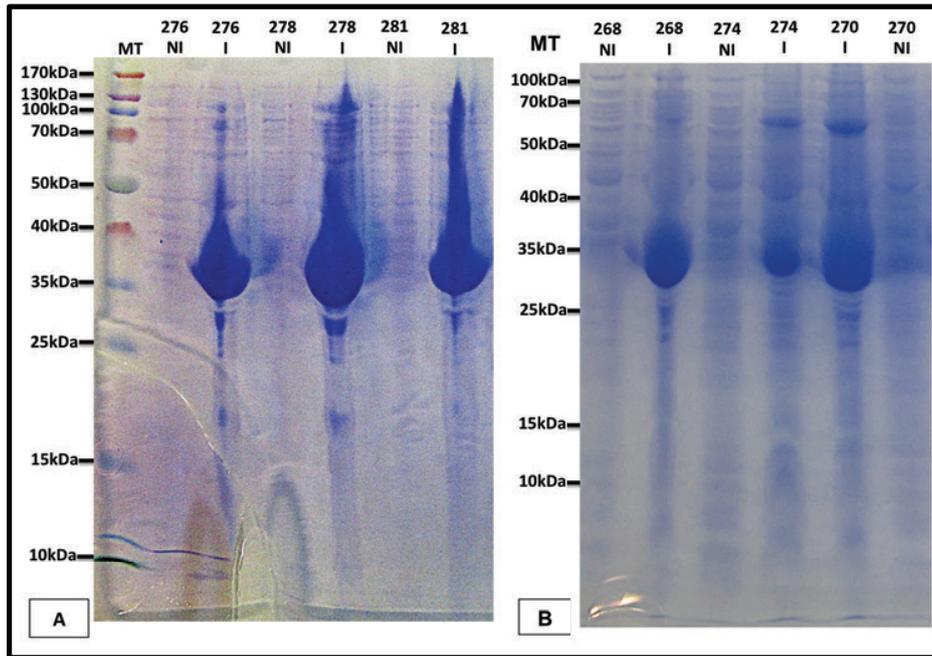


FIGURE 3.I.3 : SURPRODUCTION DES FRAGMENTS DE LA TcCRT

Analyse sur gel SDS-PAGE 12,5% en conditions réductrices et dénaturantes des extraits de protéines totales avant (**NI**) et après (**I**) induction à l'IPTG pour une quantité égale de bactéries (10 uDO). Les échantillons déposés de gauche à droite sont TcCRT Δ P Δ 276, 278 et 281 (**A**) et TcCRT Δ P Δ 268, 274 et 270 (**B**).

I.3. Purification des fragments de TcCRT

La fraction soluble des fragments de TcCRT a ensuite été purifiée en deux étapes, permettant l'élimination de la majorité des contaminants de type protéique et acide nucléique. La première étape a consisté en une chromatographie d'affinité sur une colonne de nickel; la protéine se fixe sur la colonne par l'intermédiaire de l'étiquette de six histidines, puis son élution s'effectue par compétition avec un gradient croissant d'imidazole. A une concentration de 150 mM d'imidazole, les différents fragments sont élués sous la forme d'un double pic (caractéristique de tous les fragments) (**Figure 3.I.4 (A)**). Après la réunion des fractions éluées contenant la protéine d'intérêt, l'échantillon est concentré et une dernière étape de chromatographie d'exclusion de taille permet d'éliminer le reste des contaminants ainsi que les agrégats. Les profils d'élution obtenus lors de la chromatographie d'exclusion de taille étaient similaires à celui présenté **figure 3.I.5 (A)** pour tous les fragments, à l'exception du rapport entre les pics 1 et 3, que j'aborderai dans la prochaine section. Pour chaque étape, un aliquot de la solution a été conservé afin de suivre la purification par électrophorèse en conditions dénaturantes sur gel SDS-PAGE.

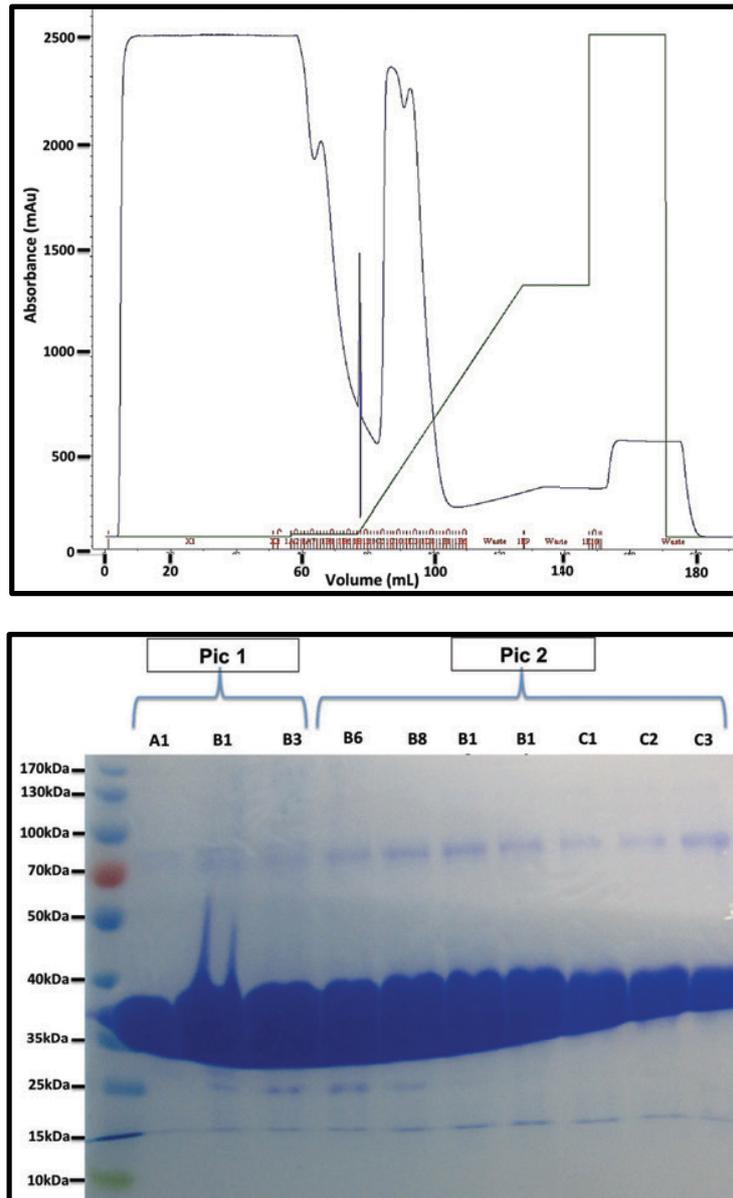


FIGURE 3.I.4 : PROFIL ET ANALYSE SDS-PAGE DE L'ETAPE DE CHROMATOGRAPHIE D'AFFINITE DE LA TcCRT Δ P Δ 281

(A) Le profil d'élution des fragments de TcCRT lors de la chromatographie d'affinité IMAC montre un double pic caractéristique élué à un volume correspondant à une concentration de 150 mM d'imidazole. (B) L'analyse sur gel SDS-PAGE montre la présence en grande quantité de TcCRT dans les deux pics, qui sont tout de même séparés par précaution.

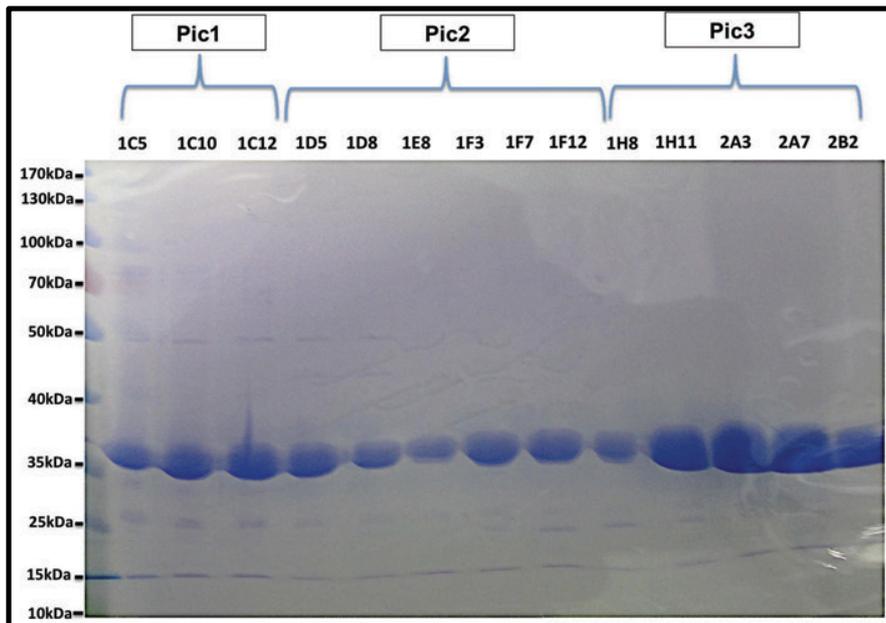
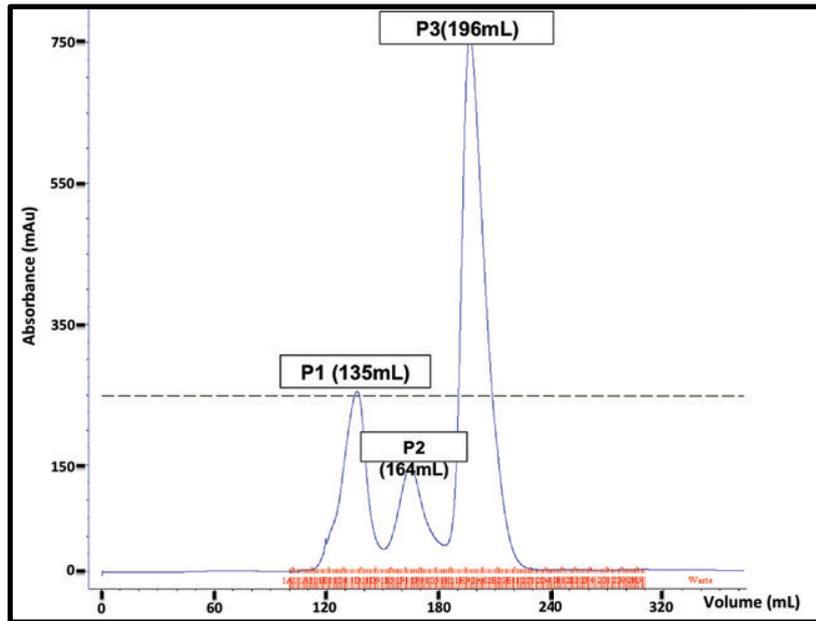


FIGURE 3.I.5 : PROFIL ET ANALYSE SDS-PAGE DE L'ETAPE DE CHROMATOGRAPHIE D'EXCLUSION DE TAILLE DE LA TcCRT Δ PA281

(A) Le profil d'élution des fragments de TcCRT lors de la chromatographie d'exclusion de taille sur une colonne de filtration sur gel S75 HiLoad 26/60 de 320 mL est systématiquement composé de 3 pics, correspondant ici à des volumes d'élution de 135 mL, 164 mL et 196 mL, respectivement. (B) L'analyse sur gel SDS-PAGE confirme la présence en grande quantité de la TcCRT très pure dans les trois pics.

I.3.a. Identification du monomère

La présence des fragments de TcCRT avec un important degré de pureté dans les différents pics de chromatographie d'exclusion de taille nous a étonnés. En effet, il est assez habituel dans ce type d'expérience d'observer plusieurs pics d'élution correspondant aux différentes formes d'oligomérisation et d'agrégation d'une protéine. Cependant, dans notre cas de figure, les volumes d'élution correspondent à des masses apparentes de 62 kDa, 33 kDa, et 16 kDa, respectivement, pour des fragments oscillant entre 30 et 35 kDa. L'observation d'un pic d'élution majeur correspondant à une masse deux fois plus petite qu'attendue était intrigante. Il était essentiel de caractériser la composition de chacun des échantillons pour les études de biologie structurale.

Nous avons saisi l'opportunité d'analyser les échantillons provenant de la purification du fragment TcCRT Δ PA Δ 281 par une expérience de SAXS. Comme je l'ai décrit dans la section Matériels et Méthodes, le SAXS permet d'obtenir une estimation de la masse moléculaire de l'échantillon en extrapolant l'intensité diffusée à un angle nul. Cette expérience a confirmé la présence du monomère dans le pic 3 de la chromatographie (**Tableau 3.I.1**), alors que les pics 1 et 2 correspondaient à des oligomères (potentiellement un trimère pour le pic 2).

| Concentration (mg/mL) | Rg (nm) | Présence d'agrégats | Dmax (nm) | Masse (kDa) |
|-----------------------|---------|---------------------|-----------|-------------|
| 1,77 | 2,44 | non | 8,52 | 30-35 |
| 2,24 | 2,46 | non | 8,63 | 30-35 |
| 6,20 | 2,51 | non | 8,77 | 32-40 |

TABLEAU 3.I.1: CARACTERISTIQUES BIOPHYSIQUES DETERMINEES PAR SAXS POUR LE FRAGMENT TcCRT Δ PA Δ 281 (PIC 3)

L'analyse SAXS des différents échantillons de chromatographie d'exclusion de taille pour le fragment TcCRT Δ PA Δ 281 a permis de déterminer différents paramètres biophysiques comme le rayon de giration (Rg), la taille maximale de la protéine (Dmax) et une estimation de la masse moléculaire, correspondant ici à celle attendue pour le monomère. La valeur mesurée à la plus haute concentration, légèrement supérieure, est possiblement due à un début d'interaction entre les particules.

Le comportement de la TcCRT, élue à une masse apparente plus petite que la masse réelle, avait déjà été observé au laboratoire lors de la purification de la calréticuline humaine. Il se pourrait que la protéine interagisse avec la matrice composée de sépharose, potentiellement par l'intermédiaire de son domaine lectine.

I.3.b. Les domaines P et C-terminaux cruciaux pour la solubilité et le repliement de la TcCRT ?

Nous avons observé l'influence de la coupure de l'hélice C-terminale sur les rendements de purification. Nous avons en effet constaté que la quantité finale de protéine obtenue pour un litre de culture bactérienne diminuait avec la longueur du fragment (**Tableau 3.I.2**).

| Fragments | Quantité moyenne obtenue pour 1L de culture |
|-------------|---|
| TcCRTFull | 30 mg |
| TcCRTΔP | 30 mg |
| TcCRTΔPΔ288 | 30 mg |
| TcCRTΔPΔ281 | 10mg |
| TcCRTΔPΔ278 | 2mg |
| TcCRTΔP 276 | 2mg |
| TcCRTΔP 274 | 1mg |
| TcCRTΔP 270 | 1,5mg |
| TcCRTΔP 268 | 1mg |

TABLEAU 3.I.2 : QUANTITE MOYENNE DE PROTEINE OBTENUE EN FIN DE PURIFICATION POUR 1 L DE CULTURE BACTERIENNE

En analysant les profils d'élution des étapes de chromatographie d'exclusion de taille, nous nous sommes également aperçus que les purifications présentant le plus faible rendement étaient associées à l'apparition d'un pic important d'agrégats et une diminution drastique du pic correspondant au monomère (**Figure 3.I.6**). Cette diminution de solubilité nous paraissait cohérente avec la suppression de nombreux acides aminés chargés.

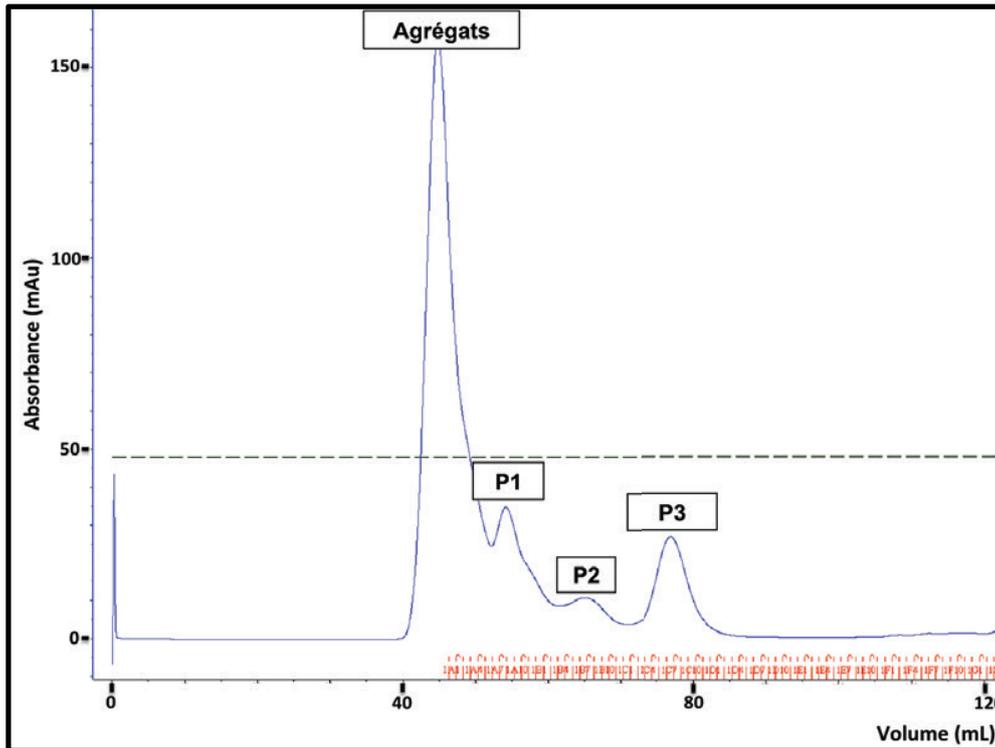


FIGURE 3.I.6 : PROFIL D'ÉLUTION DE L'ÉTAPE DE CHROMATOGRAPHIE D'EXCLUSION DE TAILLE POUR LE FRAGMENT TcCRTΔPΔ276

Le profil d'élution de chromatographie des fragments dont l'hélice C-terminale est la plus courte, comme c'est le cas ici du fragment TcCRTΔPΔ276, est similaire à celui précédemment décrit. Néanmoins, l'absorbance à 280 nm du pic correspondant au monomère (pic 3) diminue drastiquement alors qu'un important pic d'agrégat apparaît.

I.3.c. Thermofluorescence : pas uniquement un problème de solubilité

Nous avons supposé que la diminution de rendement observée lors des purifications pouvait également provenir d'une déstabilisation de la structure tridimensionnelle de la protéine. Pour vérifier cette hypothèse, nous avons observé les profils de thermofluorescence réalisés par la plateforme HTX de l'EMBL à Grenoble. Dans ces expériences, un signal de fluorescence est émis lorsque le fluorophore se fixe sur les régions hydrophobes de la protéine. Ce signal devient d'autant plus important que la protéine expose des zones hydrophobes et donc se dénature. Les résultats nous ont confirmé que la température de demi-vie (et par conséquent la stabilité) des fragments diminue avec la longueur de l'hélice C-terminale et que certains d'entre eux étaient même totalement dénaturés à température ambiante (**Figure 3.I.7**).

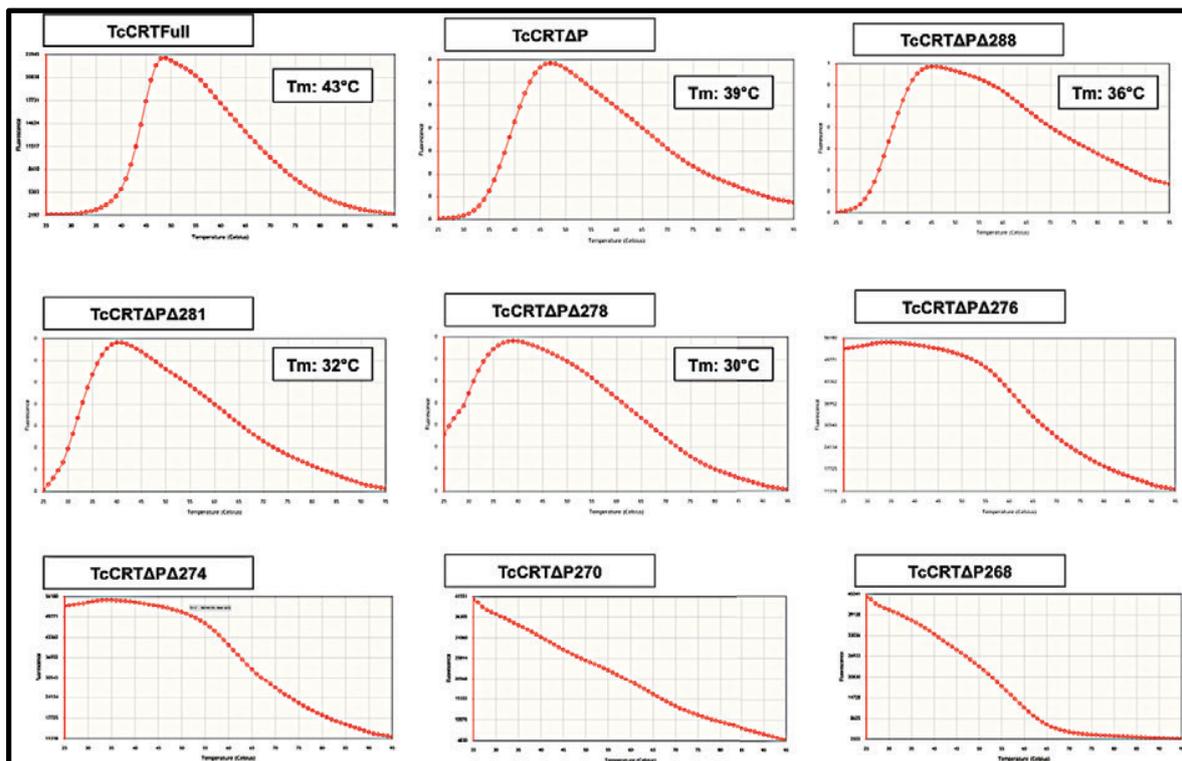


FIGURE 3.1.7 : MESURE DE LA STABILITE DES FRAGMENTS DE TcCRT PAR THERMOFLUORESCENCE

Les expériences de thermofluorescence montrent une diminution de la stabilité des fragments de TcCRT suite à la coupure du domaine P ($T_{m_{TcCRTFull}} = 43^{\circ}\text{C}$ et $T_{m_{TcCRT\Delta P}} = 39^{\circ}\text{C}$), tandis que les troncations effectuées dans l'hélice C-terminale déstabilisent encore plus la protéine jusqu'à atteindre une température de demi vie de 30°C . Enfin les fragments possédant une hélice finissant à l'acide aminé 276 sont dénaturés à température ambiante. L'intensité de fluorescence est présentée en ordonnée, tandis que l'axe des abscisses montre la température (comprise entre 25 et 95°C).

I.3.d. Dichroïsme circulaire et fluorescence : identification des fragments stables

Des expériences de dichroïsme circulaire et de fluorescence ont été effectuées pour confirmer nos différentes observations sur l'instabilité de certains fragments et pour identifier les fragments les plus stables, nécessaires aux essais de cristallisation. L'expérience de fluorescence repose sur la longueur d'onde d'émission de fluorescence des tryptophanes contenus dans la protéine. Cette dernière est directement dépendante de l'environnement du résidu. En effet un maximum de fluorescence à une longueur d'onde de 325 nm correspond à un environnement hydrophobe pour la totalité des tryptophanes, en accord avec une protéine repliée, ce qui est le cas pour les fragments TcCRTFull, ΔP , $\Delta P\Delta 288$, $\Delta P\Delta 281$ et $\Delta P\Delta 288\text{ter}$ (que je présenterai plus tard). Un maximum de fluorescence à une longueur d'onde de 350 nm correspond à une exposition des tryptophanes au solvant et suggère donc une dénaturation

probable de la protéine. Cette dénaturation touche les fragments TcCRT Δ P276 et 278 et 268 et dans une moindre mesure 270 (maximum à 340 nm) (**Figure 3.I.8**).

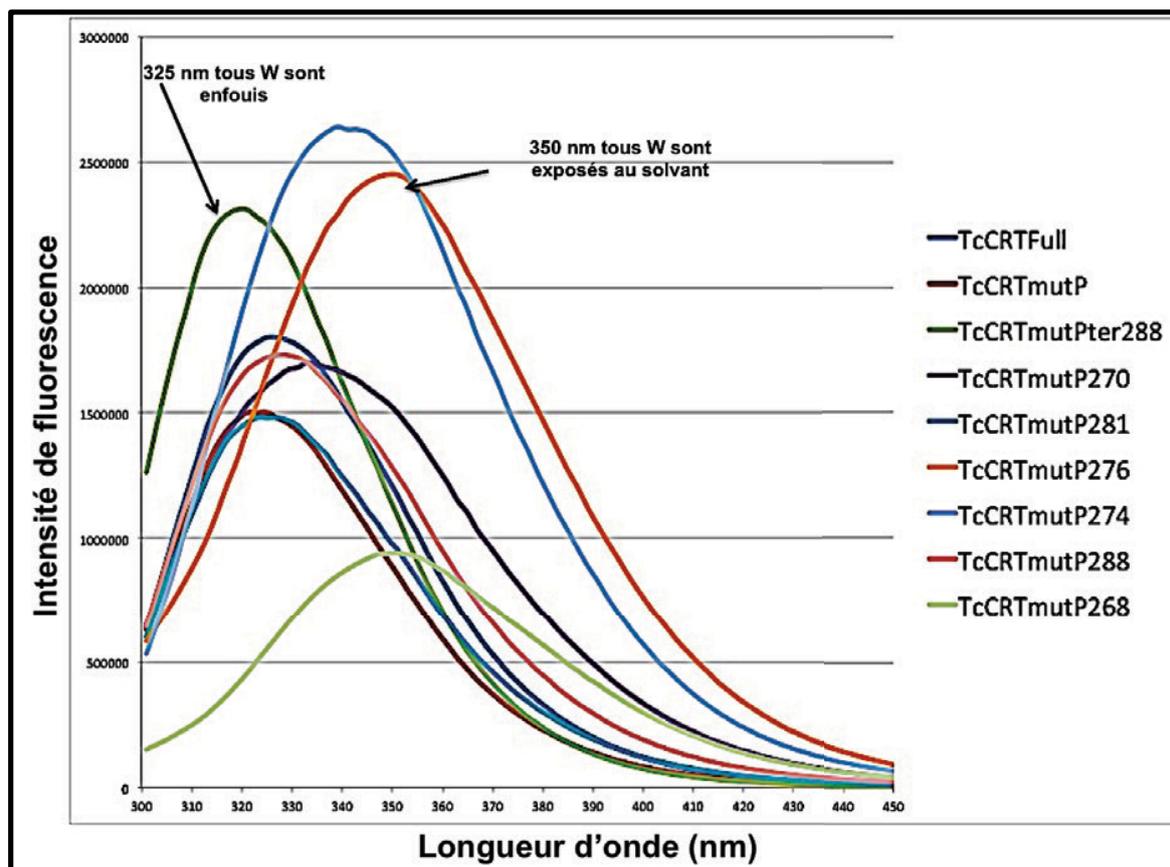


FIGURE 3.I.8 : SPECTRE DE FLUORESCENCE DES TRYPTOPHANES DES DIFFERENTS FRAGMENTS DE TcCRT

Les différents fragments de calréticuline sont soumis à un rayonnement correspondant à la longueur d'onde d'excitation des tryptophanes (280 nm) et l'intensité de fluorescence réémise est enregistrée pour toute la gamme de longueurs d'onde de 300 à 450 nm. Une émission de fluorescence maximale à une longueur d'onde de 325 nm est associée à un environnement hydrophobe pour l'ensemble des résidus W de la protéine (enfouis), tandis qu'une émission de fluorescence maximale à une longueur d'onde de 350 nm correspond à un environnement hydrophile pour ces résidus (exposés au solvant).

Le dichroïsme circulaire repose quant à lui sur la différence d'absorption de la lumière polarisée à gauche ou à droite par une molécule chirale. Dans la protéine, la liaison peptidique est responsable de la chiralité de la molécule et le signal mesuré sera spécifique de la composition en structures secondaires de la protéine comme décrit dans la section Matériels et Méthodes. La TcCRT étant composée essentiellement de feuillets β (structure secondaire participant le moins au signal) et notre difficulté à éliminer les sels dans le tampon ont conduit à l'enregistrement d'un signal de faible intensité et de qualité moyenne, dans lequel nous n'étions pas en mesure d'observer le pic spécifique aux hélices α et le pic caractéristique des

structures non repliées. Néanmoins la valeur de l'ellipticité molaire par résidu pour une longueur d'onde comprise entre 215 et 230 nm est associée à la composition en feuillets β et hélices α . Ce signal tend à disparaître pour les fragments TcCRT Δ P Δ 278, 276, 270 et 268 (Figure 3.I.9).

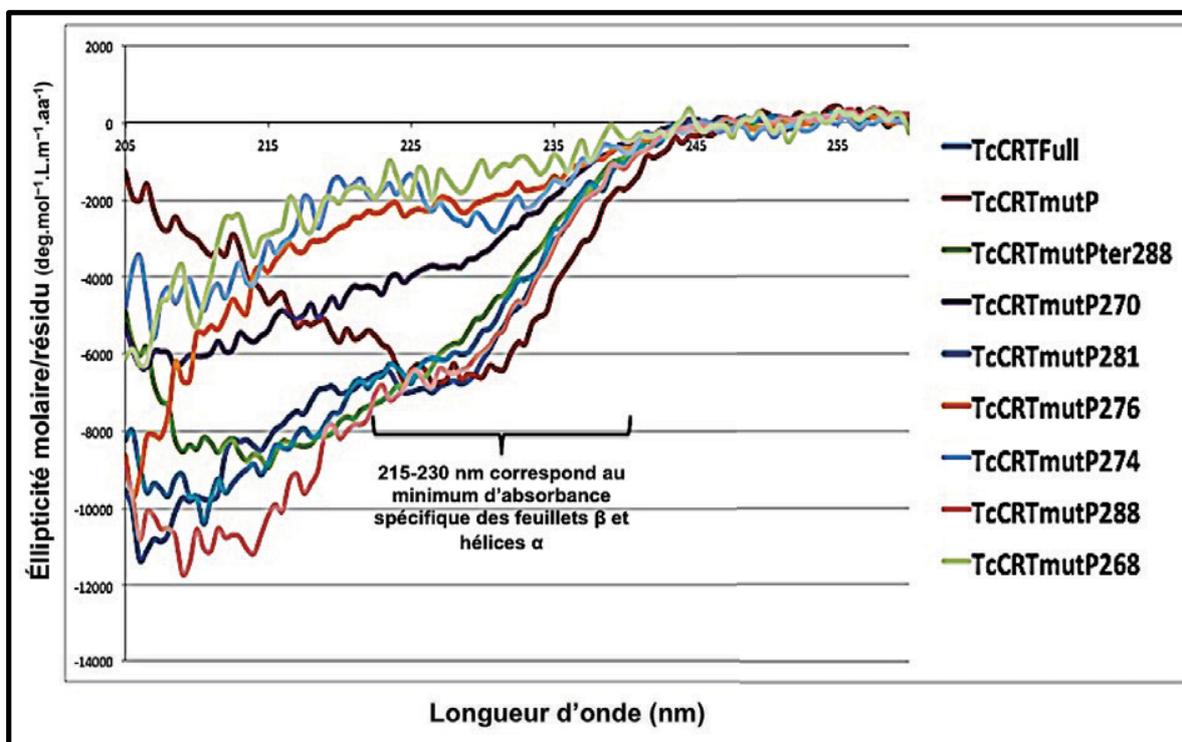


FIGURE 3.I.9 : SPECTRE DE DICHROÏSME CIRCULAIRE DES FRAGMENTS DE TcCRT

L'ellipticité molaire par résidu mesurée à partir de la différence d'absorbance de la lumière polarisée par les différents fragments de TcCRT conduit à un pic négatif pour une longueur d'onde entre 215 et 230 nm spécifique des hélices α et feuillets β . La perte du signal correspond à une diminution du contenu en structure secondaire de la protéine étudiée.

Les différentes observations effectuées lors des purifications, ainsi que les expériences de thermofluorescence, fluorescence et dichroïsme circulaire nous ont apporté des résultats cohérents entre eux, nous conduisant à regrouper nos fragments selon leur stabilité. Ainsi, les fragments TcCRTFull, TcCRT Δ P, TcCRT Δ P Δ 288 et TcCRTP Δ 281 ont été considérés comme stables et seront utilisés pour la suite des essais de cristallisation.

II. Cristaux des calréticulines de *T. cruzi* et *E. histolytica*

Environ 1200 conditions de cristallisation ont été testées pour chacun des différents fragments, identifiés comme stables par la plateforme HTX hébergée à l'EMBL de Grenoble, à différentes concentrations (couramment 5, 10 et 20 mg/mL) et à 4°C principalement. Pour chaque fragment, les pics correspondants aux pics 2 et 3 de chromatographie d'exclusion de taille ont été utilisés, soit un total d'environ 45 000 tests. Malheureusement, après un an, nous n'avons obtenu aucun cristal et même si quelques pistes ont été explorées à la main (environ 2000 essais supplémentaires), nous avons décidé de revoir notre stratégie.

II.1. Echec de la cristallisation: la calréticuline de *T. cruzi*, une protéine toujours trop flexible ?

II.1.a. Analyse SAXS : vers la dissection gagnante

La troncation initiale du domaine P a été choisie en se basant sur la structure de la calréticuline de souris (**Figure 3.I.1**). Nous nous sommes donc posé la question de savoir si le choix de cette troncation était adapté au cas de la CRT de *T. cruzi*. Pour cela nous avons utilisé les données de SAXS obtenues sur le fragment TcCRT Δ P Δ 281, qui nous avaient permis d'identifier le monomère. Ces mesures de SAXS nous ont permis de déterminer certains paramètres de la protéine, comme décrits dans la section Matériels et Méthodes. Ainsi nous avons déterminé un Dmax de 8,7 nm et un rayon de giration (Rg) de 2,5 nm. La courbe de distribution des distances (**Figure 3.II.1 (B)**) est en accord avec nos attentes, puisqu'elle correspond à une protéine globulaire possédant un domaine allongé extérieur au domaine globulaire.

De façon plus intéressante, le Kratky plot nous renseigne sur la forme globale de l'objet en solution, mais également sur sa flexibilité. En effet lorsque q tend vers l'infini, $q^2 I$ doit tendre vers 0. La divergence de la fonction par rapport à l'abscisse est associée à une importante flexibilité au sein de la protéine étudiée (**Figure 3.II.1 (A)**). De plus, nous avons en parallèle utilisé le logiciel GASBOR pour modéliser une enveloppe représentant la forme globale de la protéine. Pour effectuer cette étape de modélisation, le logiciel utilise des paramètres tels que le Rg et la Dmax pour définir des limites de distance. Au cours d'une expérience de SAXS, nous observons la conformation moyenne de toutes les particules en solution. De ce fait, si un

domaine de la protéine est flexible, la modélisation aura des difficultés à converger vers une solution unique comme nous pouvons l'observer (**Figure 3.II.1 (C)**).

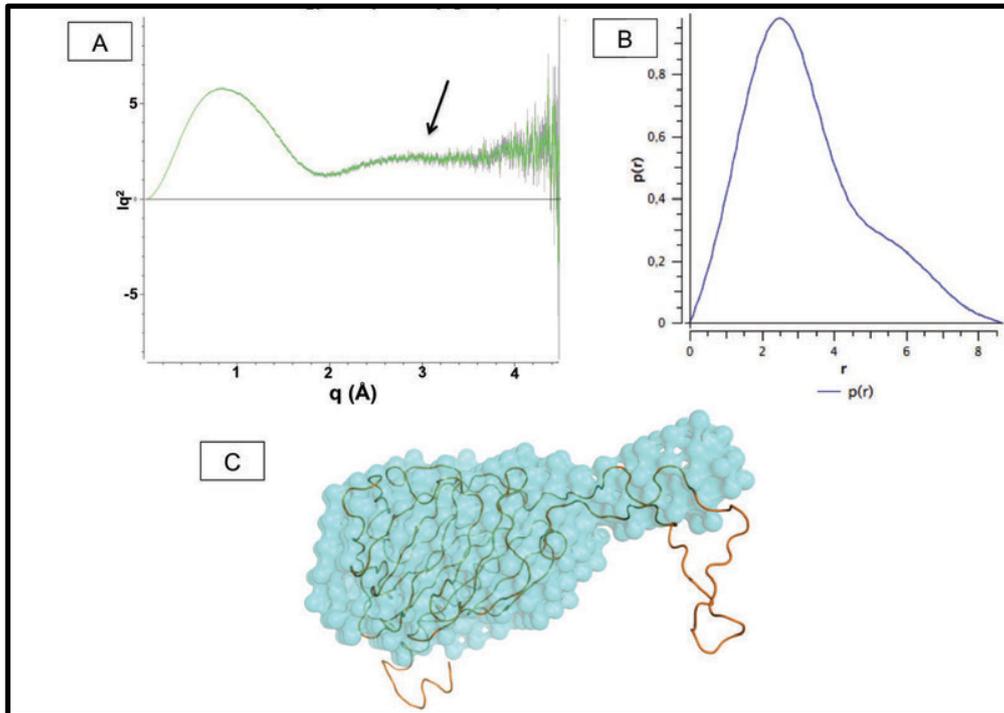


FIGURE 3.II.1 : ETUDE DE LA FLEXIBILITE DU FRAGMENT TcCRTΔPΔ281 PAR SAXS

Les expériences de SAXS menées sur le fragment TcCRTΔPΔ281 ont montré la présence d'une importante flexibilité. **(A)** Le Kratky plot présente une forme en cloche caractéristique d'une protéine plutôt globulaire, alors que la divergence de la fonction avec l'augmentation de q (flèche noire) est associée à la présence d'un domaine flexible. **(B)** La fonction de distribution des distances est représentative des distances interatomiques au sein de la protéine. Sa forme en gaussienne suivie d'un épaulement est caractéristique d'une protéine globulaire possédant un domaine allongé. **(C)** Différentes enveloppes ont été générées avec le logiciel GASBOR et moyennées. Le modèle de la TcCRT issu du serveur phyre² s'insère dans ces enveloppes, à l'exception de l'extrémité du domaine P impossible à modéliser.

Ces observations nous ont suggéré une flexibilité encore importante du domaine P présent dans nos fragments. De manière intéressante la zone de flexibilité semblait se situer au delà d'une région proposée par les calculs du logiciel el Nemo comme une seconde charnière. Nous avons donc décidé de générer un nouveau fragment nommé TcCRTΔPterΔ288, possédant le domaine P représenté dans la **figure 3.II.2** et une hélice C-terminale tronquée après le résidu 288, c'est-à-dire de la même longueur que le fragment ayant permis de résoudre la structure de la HuCRT.

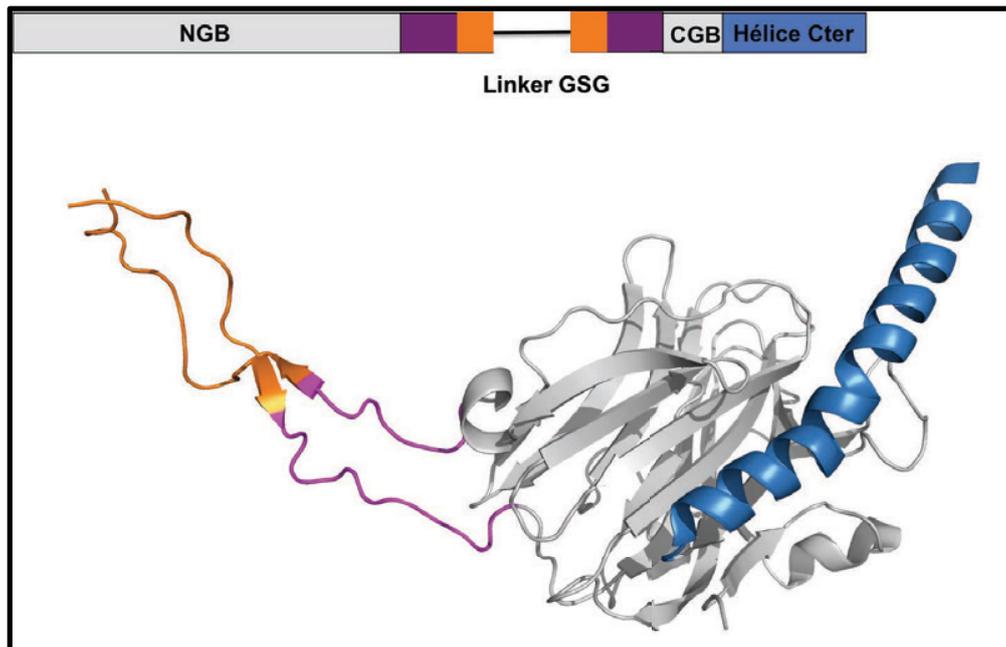


FIGURE 3.II.2 : TRONCATION SUPPLEMENTAIRE DANS LE DOMAINE P DU FRAGMENT TcCRTΔP288

Une portion supplémentaire du domaine P a été supprimée (indiquée en orange) sur la base des expériences de SAXS, dans le but d'obtenir un fragment cristallisable.

Les rendements obtenus lors de la purification de ce nouveau fragment TcCRTΔPterΔP288, sont équivalents à ceux obtenus pour la protéine entière et le fragment TcCRTΔP. Nous avons également vérifié sa stabilité par dichroïsme circulaire et spectroscopie de fluorescence (**Figures 3.I.8 et 3.I.9**). Ce fragment a à nouveau été soumis à de nombreux criblages de cristallisation afin d'évaluer le comportement de la protéine vis-à-vis d'un grand nombre de conditions. Une seule condition a permis l'apparition de cristaux, qui à première vue semblaient prendre la forme d'aiguilles très fines : 30% PEG 4000, 0,2 M acétate de sodium, 0,1 M Tris pH 8,5 à 4°C.

II.1.b. Premiers cristaux, un nouveau challenge

Les cristaux ont pu être reproduits manuellement par la technique de la goutte suspendue. Nous avons pu découvrir qu'il ne s'agissait pas d'aiguilles, mais de fines plaques organisées en bouquets que nous observions sur la tranche lors du criblage (**Figure 3.II.3**). Des tentatives d'optimisation ont été réalisées pour obtenir des cristaux uniques et plus épais (modification des rapports de gouttes, des concentrations en agent précipitant et sels, diminution de l'échange de vapeur en ajoutant un mélange de silicone/paraffine), mais sans succès.



FIGURE 3.II.3 : CRISTAUX DU FRAGMENT TcCRTΔPTERΔ288 DE LA CALRETICULINE DE *T. CRUZI*

Néanmoins, des cristaux obtenus dans la condition **35% PEG 4000, 0,2 M d'acétate de sodium, 0,1 M Tris pH 8,5** pour une concentration en protéine de 12 mg/mL présentaient de larges surfaces monocristallines nous donnant la possibilité d'enregistrer des clichés de diffraction, grâce aux lignes de lumière possédant des microfaisceaux (10 à 50 μm) comme ID23-1 et -2 et ID29. Afin d'être testés, ces cristaux ont été cryoprotégés en ajoutant 10% de glycérol à la solution de cristallisation.

Ces cristaux nous ont permis d'obtenir des jeux de données de diffraction sur la ligne de lumière de l'ESRF ID29. Ces premiers jeux de données ont été très difficiles à intégrer et indexer du fait de la présence sur chaque image du cliché de diffraction, de taches de diffraction appartenant à plusieurs cristaux. En effet, même en travaillant avec un microfaisceau, il a été très difficile d'éviter l'interaction des rayons X avec différents cristaux empilés. De plus, la très faible épaisseur des cristaux ne nous a pas permis d'obtenir des jeux de données complets dans toutes les directions de l'espace et nous avons assemblé plusieurs jeux pour obtenir des premières données complètes à une résolution de 2,9 Å. Ce cristal s'est révélé du groupe d'espace monoclinique P1 avec les paramètres de mailles suivants : **a= 79,92 Å, b=79,91 Å, c= 89 Å, $\alpha=106^\circ$, $\beta=93^\circ$, $\gamma=120^\circ$**

II.1.c. Interaction avec le glucose : amélioration des données.

Pour obtenir des cristaux de meilleure qualité, il existe différentes stratégies. Une d'entre elles consiste à tester différentes solutions de cryoprotection ; en effet, selon la nature de la solution, les cristaux peuvent ne pas la supporter et se trouver déstabilisés (par exemple par la présence des molécules de glycérol ou des contraintes mécaniques apportées par de nombreuses manipulations du cristal). Il est également possible de stabiliser la protéine par

ajout d'un ligand, afin de diminuer l'agitation et favoriser l'empilement cristallin et donc la diffraction.

Connaissant l'affinité de la TcCRT pour les protéines monoglucosylées et la capacité de certains sucres à jouer un rôle cryoprotecteur, nous avons décidé de co-cristalliser la protéine en présence de glucose. Nous avons pu constater une augmentation de la solubilité de la protéine en présence de glucose. Alors que les premiers cristaux apparaissent dans la gamme à 32% de PEG 4000, ils n'apparaissent qu'à 35% en présence de 300 mM de glucose et uniquement à 40% pour 1 M de glucose. Malheureusement la présence de glucose n'a en aucun cas modifié ni le faciès, ni la croissance des cristaux qui se présentaient toujours sous la forme d'un bouquet de fines plaques. Toutefois ces conditions nous ont permis de nous affranchir d'une solution de cryoprotection grâce à la présence du glucose à haute concentration, minimisant les manipulations du cristal.

Notre meilleur jeu de données, complet à une résolution de 2,45 Å, dont les statistiques sont présentées dans le **tableau 3.II.1** a pu être obtenu sur la ligne de lumière ID29 de l'ESRF à partir d'un cristal ayant poussé dans les conditions : **40% PEG 4000, 0,2 M acétate de sodium, 1 M glucose, 0,1 M Tris pH 8,5 à 20°C**. A nouveau, l'empilement cristallin ne possède aucune symétrie et s'organise dans un groupe d'espace P1, tandis que la présence du glucose modifie légèrement les paramètres de maille : **$a=79,3$ Å, $b=79,38$ Å, $c=85$ Å, $\alpha=95,57^\circ$, $\beta=98,7^\circ$, $\gamma=119,9^\circ$.**

| Ligne de lumière | ID29 (ESRF) |
|---|-------------------|
| Date de la collecte | 06/03/14 |
| Longueur d'onde (Å) | 0.97625 |
| Détecteur | Pilatus |
| Groupe d'espace | P1 |
| Paramètres de mailles (a, b, c) (Å) | 79.3, 79.4, 85.1 |
| Paramètres de mailles (α , β , γ) (°) | 95.6, 98.7, 119.9 |
| Résolution (haute/basse) (Å) | 2.45-82 |
| Dernière tranche de résolution (Å) | 2.45-2.60 |
| Réflexions observées | 126356 (19089) |
| Réflexions uniques | 58978 (9064) |
| Redondance | 2.14 (2.1) |
| Complétude (%) | 92 (86.6) |
| I/ σ I | 7.14(1.58) |
| Rsym (%) | 8.3 (52.2) |
| CC (1/2) | 99.3 (60.6) |
| Facteurs B Wilson | 49 |

TABLEAU 3.II.1: STATISTIQUES DES DONNEES DE DIFFRACTION D'UN JEU NATIF (AVEC GLUCOSE) DE LA PROTEINE TcCRT Δ PTER Δ 288. Les valeurs entre parenthèses correspondent à la dernière tranche de résolution.

II.2. Cristaux de la calréticuline d'*E. histolytica* (EhCRT) : utilisation des acquis?

Le succès obtenu pour résoudre la structure de la calréticuline de *T. cruzi*, nous a conduits à entreprendre la résolution de la structure de la seconde calréticuline, permettant de disséquer les fonctions portées par la CRT humaine : celle de l'amibe *E. histolytica*. Même si de nombreux tests de cristallisation ont été effectués sur la protéine entière, nous avons privilégié la stratégie utilisée pour la calréticuline de *T. cruzi*. Afin de gagner du temps, nous avons directement produit un fragment possédant une hélice C-terminale et un domaine P de la même longueur que ceux du fragment cristallisé pour la TcCRT. Ce fragment, nommé EhCRT Δ PA288, a été produit et purifié en suivant le même protocole que pour la TcCRT.

La rétention de la protéine sur la colonne nous a posé davantage de problèmes que pour la TcCRT. En effet, il était pratiquement impossible de l'éluer et nous perdions une quantité importante de protéine. Pour résoudre ce problème et ainsi diminuer les pertes de protéines dues aux interactions avec la colonne, nous avons décidé d'utiliser une colonne composée d'une matrice de silice plutôt que de sucres : la TSK G3000.

Même si ce changement améliore grandement l'efficacité de l'étape de chromatographie d'exclusion de taille (présence d'un pic d'élution majeur correspondant à la taille attendue pour le monomère), il n'abolit pas totalement les interactions entre la protéine et la colonne. Nous observons en effet un déplacement du volume d'élution au cours des injections consécutives pouvant correspondre à une dimérisation de la protéine sur la colonne (**Figure 3.II.4**).

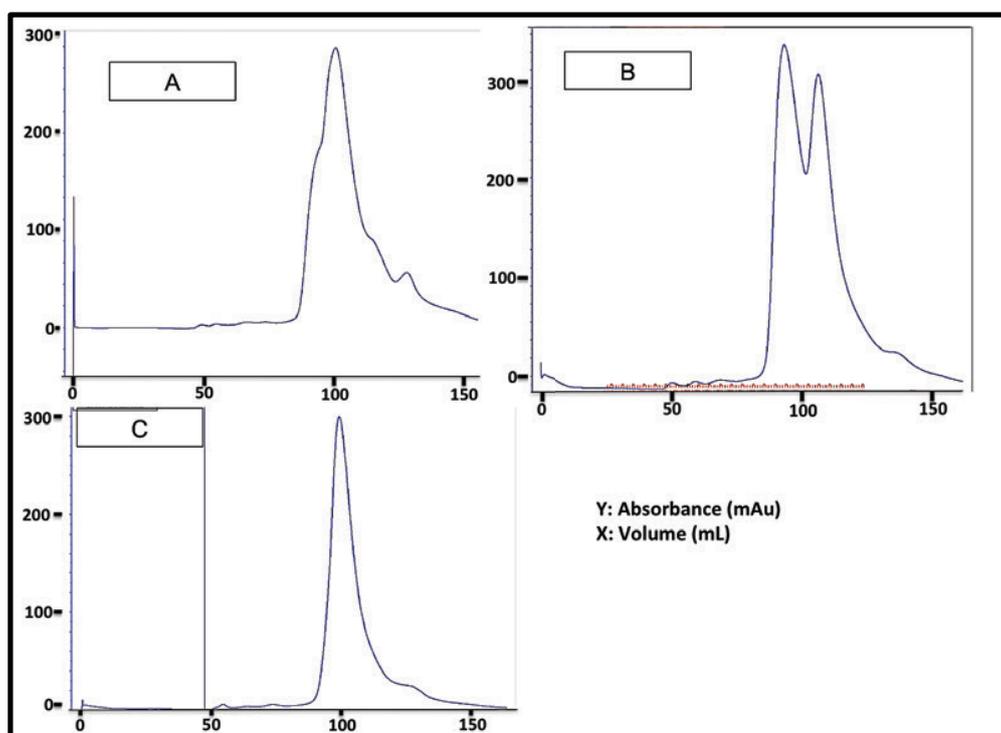


FIGURE 3.II.4 : PROFIL D'ELUTION DE L'ETAPE DE CHROMATOGRAPHIE D'EXCLUSION DE TAILLE SUR LA TSKG3000 DE LA EHCRT Δ P Δ 288

(A) Au contraire des colonnes de sépharose, le profil d'élution présente un pic unique avec un léger épaulement correspondant à un volume d'élution cohérent avec la masse moléculaire du fragment. (B) Une deuxième injection du même échantillon conduit à la formation d'un second pic, dont le sommet correspond à un volume d'élution cohérent avec un dimère. (C) Une dernière injection mène à la disparition du premier pic observé en (A), résultat d'une dimérisation totale de la protéine.

Ce fragment a ensuite été soumis à un criblage à haut débit des conditions de cristallisation à une concentration de 10 mg/mL. A nouveau, une unique condition a permis la croissance de

cristaux que nous avons pu reproduire à la main sur une gamme de 2,1 à 3,3 M de sulfate d'ammonium, 0,1 M de citrate de sodium pH 5 à 20°C. Les cristaux obtenus avaient la forme de cubes de taille variable, dont les plus gros sont représentés **figure 3.II.5**. De manière assez surprenante, deux formes cristallines étaient régulièrement obtenues dans les mêmes conditions, voire de la même goutte. Un cas particulier d'un seul cristal incluant les deux empilements cristallins a également été rencontré. Cette réorganisation de la protéine dans le cristal peut expliquer l'hétérogénéité présente dans les cristaux. En effet, selon la région du cristal centrée, la diffraction maximale pouvait atteindre 2 Å de résolution contre 20 Å à une autre position du même cristal.

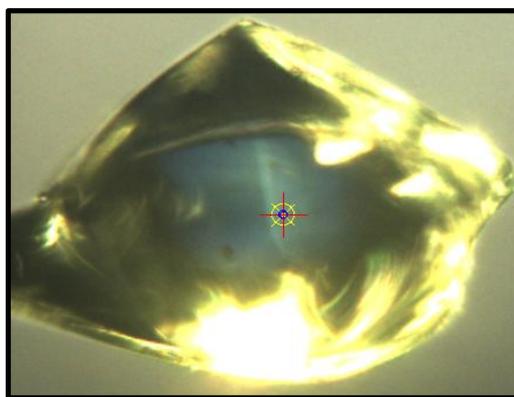


FIGURE 3.II.5 : EXEMPLE D'UN CRISTAL OBTENU POUR LE FRAGMENT EHCRTΔPΔ288

Les cristaux obtenus pour la EhCRT avaient l'aspect de cubes pouvant atteindre une taille de 400 μm sur 300 μm.

Les deux meilleurs jeux de données de diffraction pour chacune des formes cristallines ont pu être enregistrés sur les lignes de lumière ID23-1 et ED23-2, avec des cristaux ayant poussé dans les conditions **de sulfate ammonium 2,7 M, citrate de sodium 0,1 M pH 5 à 20°C** (cryoprotégé par une solution de sulfate d'ammonium 3 M, citrate de sodium 0,1 M pH 5, 20% Glycérol) et **de sulfate ammonium 2,6 M, O-phosphosérine 20 mM, citrate de sodium 0,1 M pH 5 à 20°C** (cryoprotégé par une solution sulfate d'ammonium de 3 M, O-phosphosérine 20mM, citrate de sodium 0,1 M, pH 5, 20% Glycérol).

Pour la première condition, l'empilement cristallin s'organise autour de deux axes hélicoïdaux 4_2 et 2_1 et d'un axe de symétrie d'ordre 2 dans un groupe d'espace tétragonale ($P4_22_12$) et nous a permis d'obtenir un jeu de données complet jusqu'à une résolution de 2.9 Å. Dans la seconde forme cristalline, les unités asymétriques sont reliées entre elles par trois axes

hélicoïdaux 2_1 dans un groupe d'espace $P2_12_12_1$. Cette seconde forme nous a permis d'enregistrer un jeu de données complet à une résolution de 2.15 Å (**Tableau 3.II.2**).

| | Forme cristalline 1 ($P4_22_12_1$) | Forme cristalline 2 ($P2_12_12_1$) |
|---|---|---|
| Ligne de lumière | ID23-2 (ESRF) | ID23-1 (ESRF) |
| Date de la collecte | 07/10/14 | 07/12/14 |
| Longueur d'onde (Å) | 0.87260 | 0.97625 |
| Détecteur | Pilatus | Pilatus |
| Groupe d'espace | $P4_22_12_1$ | $P2_12_12_1$ |
| Paramètres de mailles (a, b, c) (Å) | 149.3, 149.3, 117.0 | 74.4, 143.4, 171.6 |
| Paramètres de mailles (α , β , γ) (°) | 90, 90, 90 | 90, 90, 90 |
| Résolution (haute/basse) (Å) | 2.9-47.2 | 2.15-43 |
| Dernière tranche de résolution (Å) | 2.9-3.07 | 2.15-2.20 |
| Réflexions observées | 263346 (42163) | 354918 (22073) |
| Réflexions uniques | 29847 (4641) | 96632 (6384) |
| Redondance | 8.82 (9.08) | 16.07 (3.48) |
| Complétude (%) | 99.5 (97.8) | 99.2 (99.1) |
| I/ σ | 16.08 (1.82) | 13.40 (1.57) |
| Rsym (%) | 10.3 (90.6) | 5.9 (88.5) |
| CC (1/2) | 99.8 (70.1) | 99.8 (53.6) |
| Facteurs B Wilson | 72.4 | 51.5 |

TABLEAU 3.II.2 : STATISTIQUES DES DONNEES DE DIFFRACTION DES DEUX JEUX NATIFS DE LA PROTEINE EHCRT Δ PA288. Les valeurs entre parenthèses correspondent à la dernière tranche de résolution

Nous avons également testé l'ajout du glucose comme ligand pour la EhCRT, mais cela a diminué la croissance des cristaux et défavorisé la diffraction, potentiellement du fait d'une déstabilisation de l'empilement cristallin.

III. Comparaison structurale des calréticulines de parasites et de mammifères

La résolution des structures par remplacement moléculaire des calréticulines de *T. cruzi* et *E. histolytica*, dont les statistiques des modèles sont reportées dans le **tableau 3.III.1**, nous a permis de confirmer la présence d'une organisation modulaire similaire entre les espèces (domaine P, domaine globulaire et hélice C-terminale) et également de mettre en évidence des motifs consensus et certaines divergences que je vais décrire à un niveau moléculaire.

| | TcCRT | EhCRT (P4 ₂ 2 ₁ 2) | EhCRT (P2 ₁ 2 ₁ 2 ₁) |
|-----------------------------------|--------------------|--|--|
| Résolution (Å) | 2.45 | 2.9 | 2.15 |
| Rfact (%) | 22.29 | 18.80 | 18.8 |
| Rfree (%) | 25.66 | 21.56 | 21.78 |
| Déviations liaisons (Å) | 0.004 | 0.003 | 0.009 |
| Déviations angles (°) | 0.837 ₁ | 0.889 | 1.18 |
| Ramachandran outliers (%) | 0.1 (<0.2) | 0.4 (<0.2) | 0.3 (<0.2) |
| Ramachandran favorisés (%) | 97 (>98) | 94 (>98) | 95.2 (>98) |
| Rotamères outliers (%) | 0.6 (1) | 0.6 (1) | 1.4 (1) |
| Clashscore | 4.71 | 6.91 | 1.52 |
| Cβ outliers | 0 (0) | 0 (0) | 0 (0) |
| Facteur B moyen (Å ²) | 48 | 88.7 | 53.3 |
| Score Molprobit | 1.42 | 1.8 | 1.34 |

TABLEAU 3.III.1: STATISTIQUES D’AFFINEMENT DES STRUCTURES DES FRAGMENTS TcCRTΔP_{TER}Δ288 ET EhCRTΔP_AΔ288, RESOLUES PAR REMPLACEMENT MOLECULAIRE.

III.1. Organisation des TcCRT et EhCRT dans le cristal

III.1.a Empilement de la TcCRT

L'unité asymétrique contient six monomères organisés en deux trimères. Les monomères interagissent *via* le domaine P et s'organisent autour d'un pseudo axe de rotation d'ordre 3 au centre des domaines P. A l'opposé, les hélices α C-terminales forment une surface plane et sont impliquées dans une interaction tête-bêche entre les deux trimères (**Figure 3.III.1 A**). L'interaction entre les deux trimères est médiée par une seule liaison ionique par hélice α (3 interactions maintiennent donc deux trimères ensemble) entre les résidus K360 d'une hélice et

D58 dans une boucle à proximité (**Figure 3.III.1 B**). Le faible nombre d'interactions stabilisantes peut expliquer le défaut de croissance des cristaux dans une dimension.

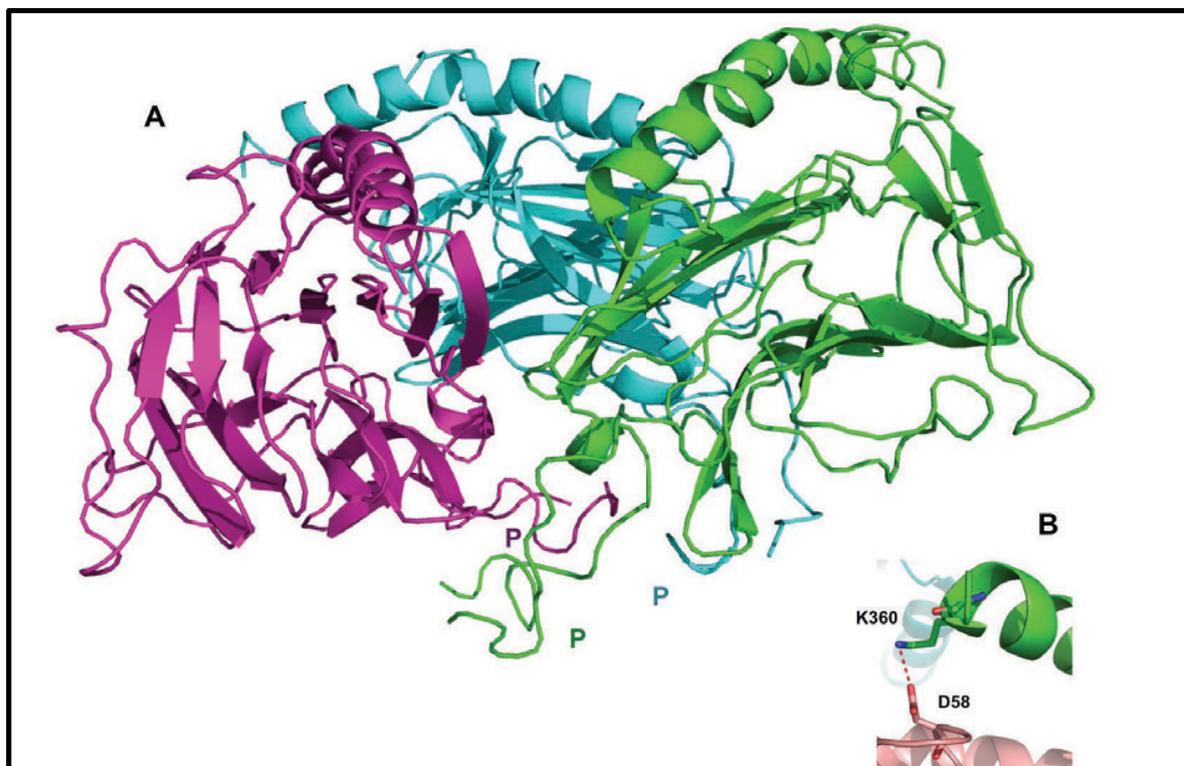


FIGURE 3.III.1 : ORGANISATION DES MONOMERES DE TcCRT DANS L'UNITE ASYMETRIQUE
(A) Les monomères interagissent *via* leur domaine P pour former deux trimères autour d'un pseudo axe de rotation d'ordre 3. A l'opposé, la surface plane, formée par les hélices α , permet l'interaction entre deux trimères. **(B)** Un maximum de deux chaînes latérales par hélice est impliqué dans ces interactions expliquant la faible croissance du cristal dans une dimension.

III.1.b Empilement de la EhCRT

Dans le groupe d'espace $P4_22_12$, l'unité asymétrique contient deux monomères qui interagissent tête-bêche via une stabilisation du domaine P d'un monomère par le domaine lectine de l'autre (**Figure 3.III.2 (A)**). L'interaction des trois monomères constituant l'unité asymétrique du groupe d'espace $P2_12_12_1$ est plus complexe. Le même type d'interaction entre le domaine P de la chaîne A et le domaine lectine de la chaîne B se produit, le domaine P de la chaîne C se replie sur son propre domaine lectine (**Figure 3.III.2 (B)**). J'affinerai ces observations à un niveau moléculaire dans la prochaine section. Nous pouvons observer quatre conformations différentes du domaine P ; cette flexibilité est certainement liée à la très grande quantité de solvant présente dans les cristaux (75%).

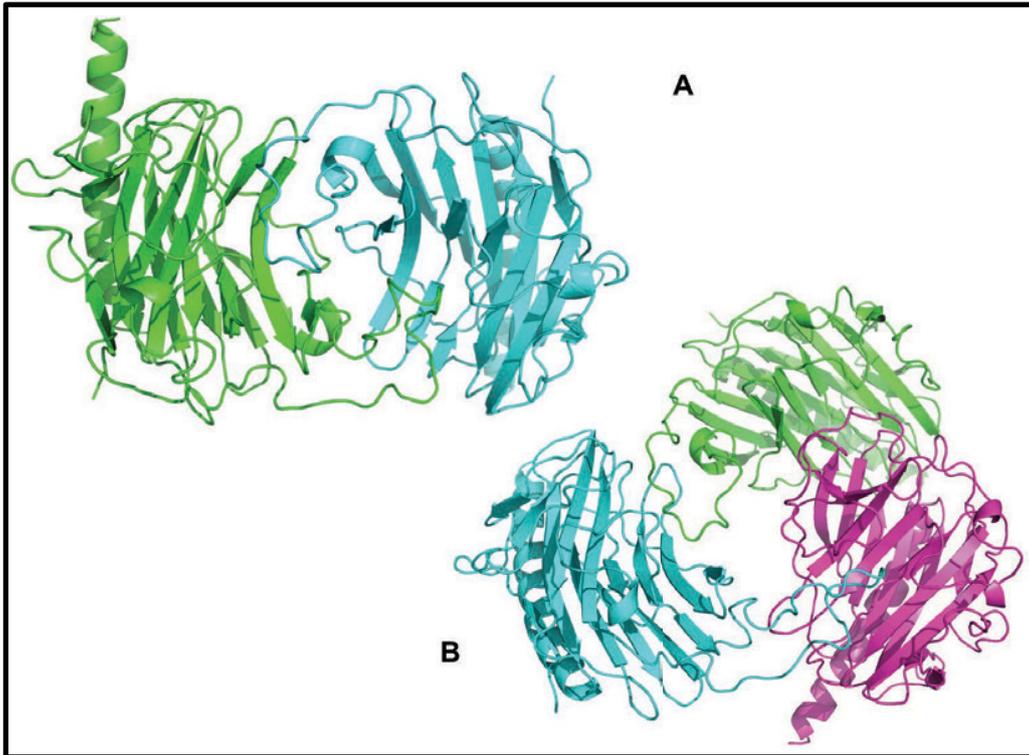


FIGURE 3.III.2 : ORGANISATION DES MONOMERES DE EHCRT DANS L'UNITE ASYMETRIQUE DES DIFFERENTS GROUPES D'ESPACE

(A) Dans la première forme cristalline ($P4_21_2$), deux monomères sont stabilisés tête-bêche par des interactions entre le domaine P et le domaine lectine. (B) Dans la seconde forme cristalline ($P2_12_12_1$) deux monomères sont stabilisés par des interactions entre le domaine P et le domaine lectine (vert et bleu), alors que c'est le domaine globulaire du 3^{ème} monomère (magenta) à la place de son domaine P qui intervient dans l'interaction avec le domaine lectine de la chaîne en vert.

III.2. Comparaison des domaines globulaires

La séquence des fragments recombinants des calréticulines de *T. cruzi* (TcCRT) et *E. histolytica* (EhCRT) présente entre 40 et 45% d'identité et 60% de similarité avec les CRTs humaine et de souris au niveau du domaine globulaire. Ce domaine inclut le site fonctionnel lectine hautement conservé, mais également un amas de petites différences que je vais décrire (Figure 3.III.3).



FIGURE 3.III.3 : ALIGNEMENT STRUCTURAL DES FRAGMENTS DE CALRETICULINES

L’alignement structural des fragments de calréticulines ayant permis la résolution des structures montre une forte homologie dans le domaine globulaire. Hormis une différence dans la nature des acides aminés (45% d’identité), les calréticulines diffèrent par de nombreuses insertions et délétions (indiquées par des losanges), dont cinq importantes entraînent un changement de conformation des structures secondaires (sites 1 à 5). Les résidus hautement conservés dans le domaine lectine sont indiqués par la lettre 1, tandis que la jonction permettant l’insertion du domaine P est représentée en bleue et indiquée par la lettre j. Les insertions correspondant au domaine P sont liées aux différences de construction des fragments pour obtenir la structure et non à une insertion réelle dans la séquence et le linker (g) est représenté en orange.

III.2.a. Une accumulation de différences

Malgré une forte ressemblance du domaine globulaire (déviations moyennes des Cα de 1 Å sur une superposition de 223 résidus), il existe un amas de petites différences entre les calréticulines des différentes espèces. Nous avons relevé cinq petites insertions/délétions d’acides aminés, situées principalement dans des boucles, et ayant la particularité d’être toutes exposées sur la même face de la protéine, à l’opposé du site lectine. Ces différences, qui incluent un changement local de structure secondaire, pourraient intervenir dans les interactions avec les ligands des CRTs et notamment avoir permis au cours de l’évolution l’acquisition de nouvelles fonctions chez les mammifères (**Figure 3.III.4**).

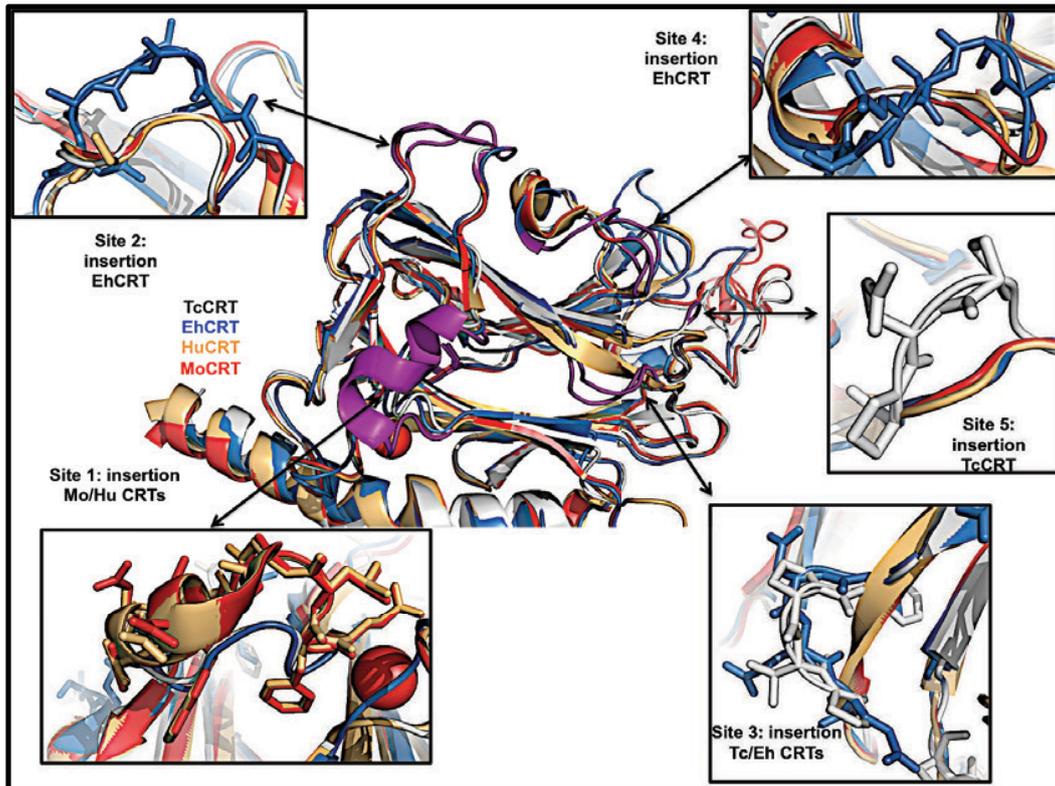


FIGURE 3.III.4 : INSERTIONS ET DELETIONS MAJEURES DANS LES FRAGMENTS DE CALRETICULINES

La superposition des structures des fragments de calréticulines de souris (**MoCRT en rouge**), de *T. cruzi* (**TcCRT en gris**), humaine (**HuCRT en orange**) et d'*E. histolytica* (**EhCRT en bleu**) montre une forte homologie du domaine globulaire incluant cinq différences principales. Ces différences, caractérisées par des insertions et délétions, sont représentées **en violet** et se situent principalement dans des boucles. La majorité de ces insertions sont présentes dans les calréticulines de parasite. En effet, les sites 2 et 4 correspondent à des insertions uniquement présentes dans la EhCRT, le site 5 à une insertion exclusive à la TcCRT, alors que le site 3 présente une insertion commune aux deux parasites. Néanmoins, une insertion importante de quatre acides aminés est présente chez les mammifères (Mo et HuCRT), à proximité du site de fixation du **calcium (sphère rouge)**; et entraîne la réorganisation d'une boucle en hélice α à l'extrémité N-terminale.

III.2.b. Site lectine et fixation du glucose

Les domaines lectines des calréticulines de mammifères et de la calnexine sont très fortement conservés (**Figure 1.II.5**), comme je l'ai présenté dans mon introduction (section II.2.d page 27) (**Chouquet et al., 2011; Kozlov et al., 2010a; Schrag et al., 2001**). Ici, la résolution des structures des fragments de calréticuline de *T. cruzi* et *E. histolytica*, deux espèces phylogénétiquement éloignées, a confirmé l'importance de ce domaine par son maintien au cours de l'évolution, pour assurer la fonction chaperonne (**Figure 3.III.5**). Ce site lectine, présent à la surface du feuillet β concave et partagé avec la calnexine, contient 19 résidus

identiques dont la conformation des chaînes latérales est strictement conservée (**Figures 3.III.3 et 3.III.5**).

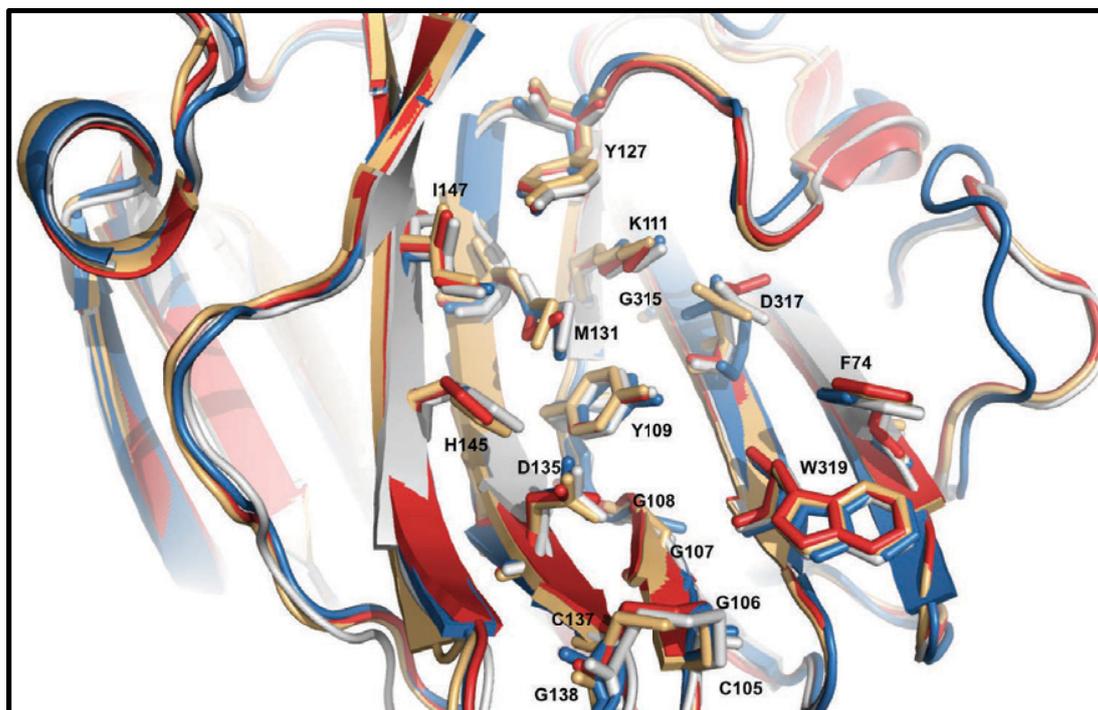


FIGURE 3.III.5 : CONSERVATION DU SITE LECTINE DANS LES CALRÉTICULINES

La superposition des calréticulines humaine (**orange**), de souris (**rouge**), d'*E.histolytica* (**bleu**) et de *T. cruzi* (**gris**) montre une importante conservation du site lectine et notamment de l'orientation des chaînes latérales des 17 résidus indiqués sur la figure. Les mutations de Y109, M131 ou K111 altèrent la fonction lectine. **La numérotation des résidus publiée pour les structures des HuCRT et MoCRT a été utilisée (Chouquet et al., 2011; Kozlov et al., 2010a).**

Au sein de ce site lectine, le site de fixation du tétrasaccharide terminal des protéines néosynthétisées a été déterminé (**Kozlov et al., 2010a**). Lors de la résolution de la structure du fragment de TcCRT co-cristallisé avec du glucose, nous avons été en mesure d'observer une molécule de glucose dans le site lectine pour deux des six monomères présents dans l'unité asymétrique (remplacé par un acétate dans les autres). Ici la fixation est légèrement différente de ce qui avait été décrit pour la calnexine et la calréticuline de souris (**Kozlov et al., 2010a; Schrag et al., 2001**). La première position du glucose est perpendiculaire à la surface concave du site lectine, potentiellement due à l'interaction du domaine P d'un autre monomère à proximité que j'aborderai dans une prochaine partie. La seconde position, bien que parallèle au site lectine, est légèrement décalée et enfouie moins profondément dans le site lectine. De plus, les molécules de glucose semblent avoir pivoté de 45°, bien que cette observation puisse être due à une mauvaise interprétation de la carte. En effet, la plus grande

distance entre la molécule de glucose et le site lectine et donc sa mauvaise stabilisation peuvent être à l'origine de la mauvaise qualité de la carte de densité. Une compétition avec les ions acétate observés dans les autres monomères pourrait en être l'explication (**Figure 3.III.6 et 3.III.9**).

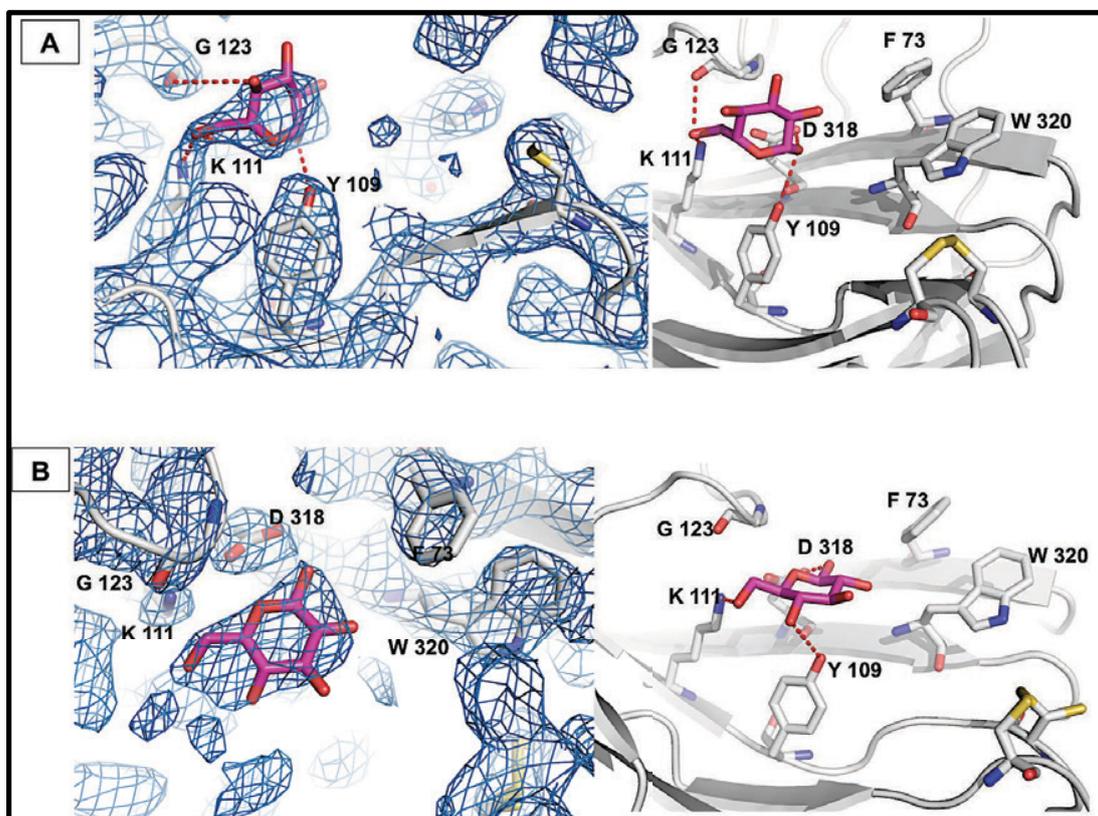


FIGURE 3.III.6 : INTERACTION DU GLUCOSE DANS LE SITE LECTINE DE LA TcCRT

L'interaction du glucose dans le site lectine de la TcCRT est stabilisée par de nombreuses liaisons hydrogènes. Les oxygènes des groupements hydroxyles O1, O5 et O6 du glucose interagissent avec les chaînes latérales des résidus Y109, K111, G123 et D318. Deux positions du glucose ont été observées ; la première proche de la position consensus présente le glucose perpendiculaire au site lectine (**A**), alors que la seconde montre le glucose parallèlement (**B**).

Lors de la résolution de la structure de la EhCRT dans le groupe d'espace P4₂2₁2, la carte Fo-Fc montrait un important pic positif, qui ne correspondait à aucun des ligands introduits dans le réservoir. Le logiciel Phenix permet de cribler les 500 ligands les plus souvent rencontrés dans la Protein Data Bank (PDB). De cette matière 170 molécules ont été sélectionnées comme des candidats potentiels pouvant correspondre à la densité observée. Les cinq molécules possédant la plus grande corrélation sont présentées **figure 3.III.7**. Parmi ces cinq candidats (N-acétyl D-allosamine, arginine, α -D-glucose, β -D-glucose et acide orotique), le glucose a été sélectionné car il s'agissait de la solution la plus probable physiologiquement.

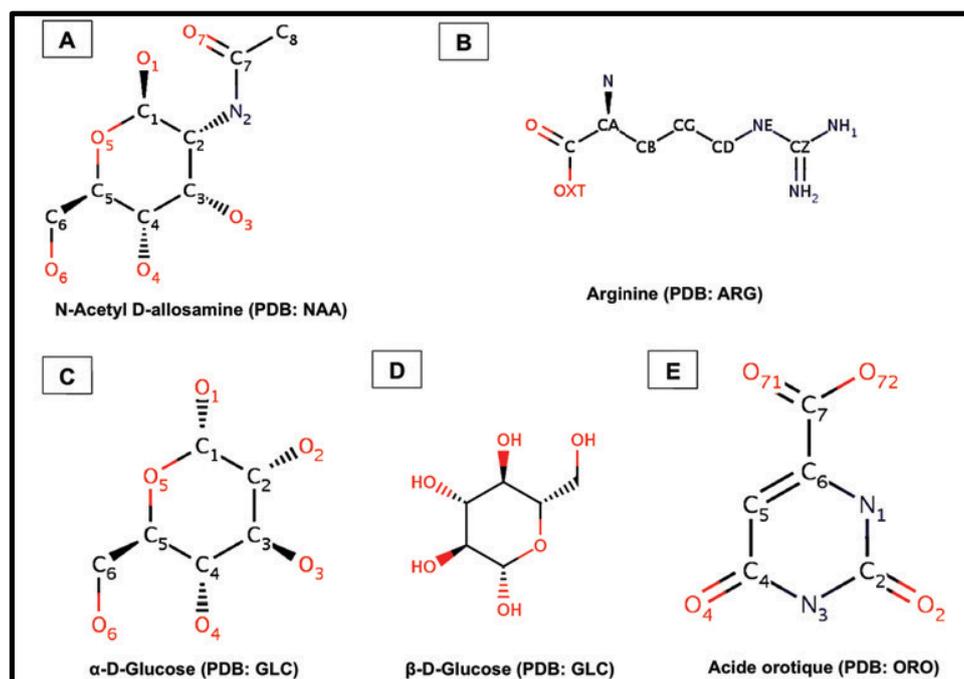


FIGURE 3.III.7 : IDENTIFICATION DU LIGAND PRESENT DANS LE DOMAINE LECTINE DE LA EhCRT

Le logiciel Phenix a permis d'établir une liste de ligands potentiels pouvant correspondre à la densité électronique observée dans les cristaux de EhCRT. Ces ligands sont classés en fonction d'un coefficient d'ajustement à la densité électronique. Les cinq meilleurs scores : le NAA (score 0,834) (A), le résidu Arg (score 0,83) (B), l' α -D-Glucose et le β -D-Glucose (score 0,82) (C et D) et l'ORO (score 0,81) (E).

Il est alors possible que le glucose, piégé en sandwich entre deux monomères dans l'empilement cristallin, puisse être issu du milieu de culture lors de la lyse des cellules. Les cartes 2Fo-Fc et de différences Fo-Fc sont représentées **figure 3.III.8**. L'interaction avec le glucose est retrouvée dans de nombreux cristaux de la EhCRT de manière reproductible. Contrairement à la TcCRT, la position du glucose ici se rapproche de celle observée dans la MoCRT avec le tétrasaccharide, en se positionnant à plat, parallèle au domaine lectine, bien qu'un retournement de 180° soit observé. (**Figure 3.III.9 et 10**).

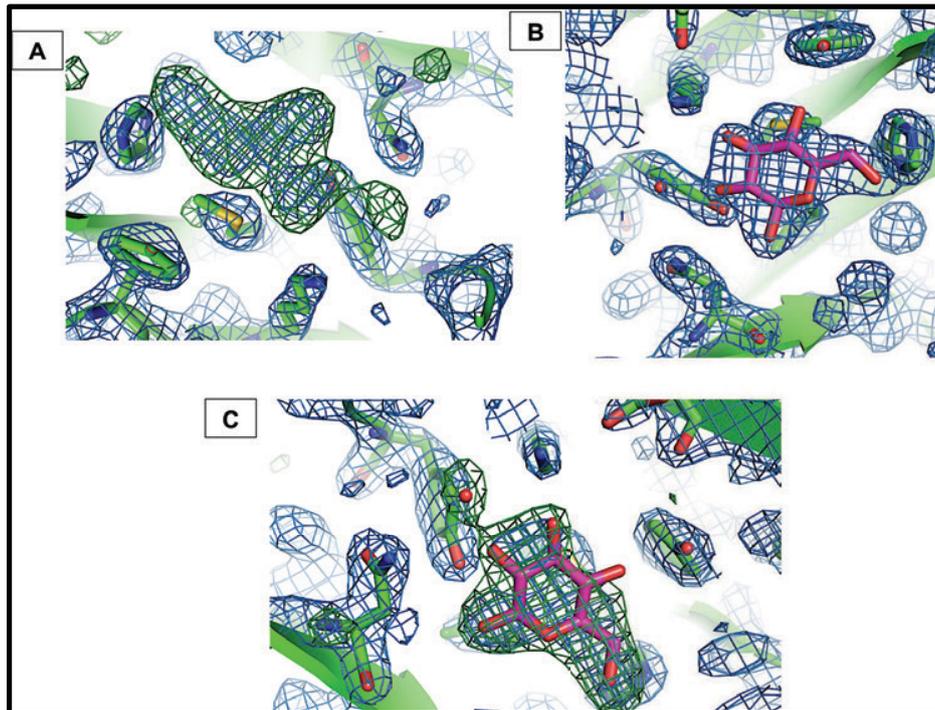


FIGURE 3.III.8 : CARTES DE DENSITE DE LA MOLECULE DE GLUCOSE DANS LE SITE LECTINE DE LA EHCRT

(A) La carte de différence Fo-Fc (verte) présente un pic positif ne correspondant pas à une molécule introduite dans la solution de cristallisation au milieu du domaine lectine. (B) Le programme Phenix Ligand Identification en criblant les nombreux ligands de la PDB a fait correspondre le glucose à la densité avec un score très élevé. (C) Le glucose dans la carte 2Fo-Fc du modèle final (bleue), superposé à la Fo-Fc initiale. Les cartes 2Fo-Fc et Fo-Fc sont contourées à 2 et 3 σ respectivement.

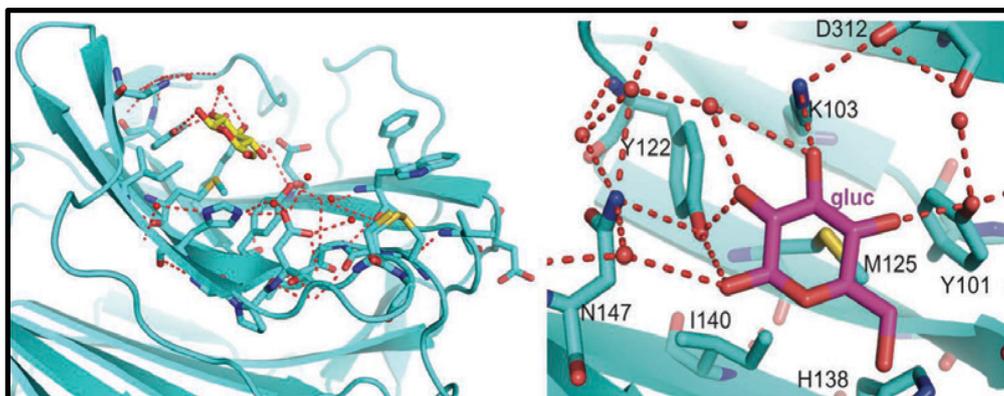


FIGURE 3.III.9 : SITE DE FIXATION DU GLUCOSE DANS LA EHCRT

Le glucose se positionne parallèlement au site lectine (gauche). Dans ce cas les nombreux groupements hydroxyles du glucose sont stabilisés par des liaisons hydrogène directes avec les résidus Y101, K103 et Y122. De plus, la molécule de glucose est stabilisée par un réseau de liaisons hydrogène relié par des molécules d'eau, tandis que les résidus M125 et I140 stabilisent le cycle par des contacts de type Van der Waals.

Bien que la position du glucose diffère légèrement par rapport à celle observée dans le contexte du tétrasaccharide complet, de nombreux résidus communs sont impliqués dans l'interaction. Le cycle du glucose se couche dans les deux cas au dessus du sillon hydrophobe entre les chaînes latérales des méthionine et isoleucine conservées. De plus, de nombreuses liaisons hydrogène impliquent les chaînes latérales d'une lysine conservée (K103 pour la EhCRT, K111 pour la MoCRT et la TcCRT) et au moins une des deux tyrosines (Y101 et Y122 pour la EhCRT, Y128 pour la MoCRT et Y109 pour la TcCRT) (**Figure 3.III.9 et 3.III.10**).

Ces observations sont cohérentes avec la description d'un site de fixation du glucose impliquant les résidus conservés K, M, I et les deux Y. En effet les mutations de ces résidus diminuent la fixation du carbohydrate dans la CRT et CNX de mammifère (**Kozlov et al., 2010a**). Ces observations soulignent une certaine versatilité dans le mécanisme de reconnaissance du glucose.

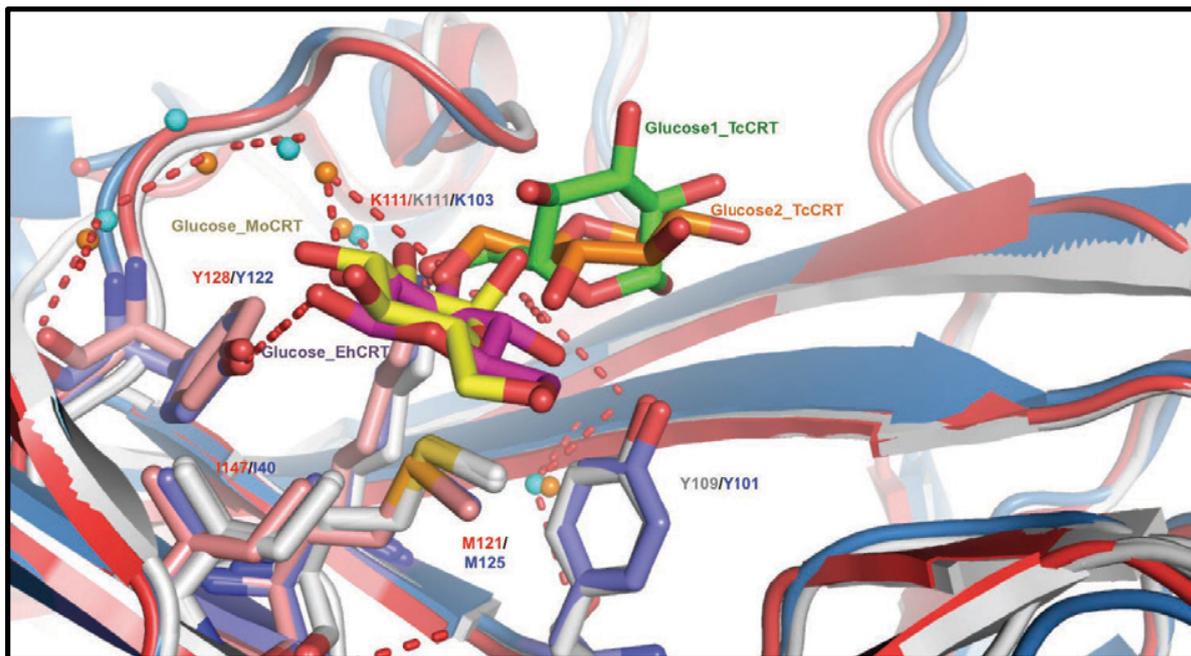


FIGURE 3.III.10 : SITE DE FIXATION DU GLUCOSE CONSERVE ENTRE LES CALRETICULINES

La superposition des calréticulines de souris (**rouge**), de *E. histolytica* (**bleu**) et *T. cruzi* (**gris**) montre une différence de positionnement des glucoses. En comparaison du glucose terminal du tétrasaccharide (**en jaune**), celui observé dans la EhCRT (**violet**) bascule de 180°, alors que les glucoses observés dans la TcCRT tournent de 45° et sont moins enfouis dans le site lectine. Néanmoins un site de fixation du glucose impliquant les résidus Y (128 et 122 et/ou 109 et 101), K (111 et 103) et les résidus méthionine et isoleucine formant un sillon hydrophobe semble se dessiner.

III.2.c. Site de fixation du calcium et du chlore

Contrairement à la TcCRT, la EhCRT possède un site de fixation du calcium identique à ceux observés dans les Mo et HuCRT, comme abordé dans l'introduction (**Figure 3.III.11**).

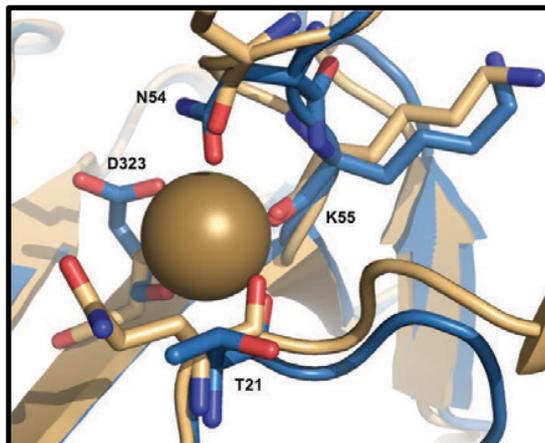


FIGURE 3.III.11 : SITE DE FIXATION DU CALCIUM DANS LA EHCRT

Le site de fixation du calcium est composé des chaînes latérales des résidus D323 et N54, ainsi que des carbonyles des chaînes principales des résidus T21 et K55. Les différents résidus impliqués dans la fixation du calcium sont conservés avec ceux jouant le même rôle dans les CRTs humaine et de souris représentées ici en bleu (D326, Q26, K62, K64).

De manière plus surprenante, un ion chlorure a été observé au voisinage du site lectine, coordonné par la chaîne latérale de l'histidine 138, un des résidus conservés du site lectine et par la chaîne principale de la valine 139 (EhCRT) ou de l'isoleucine 139 (TcCRT). L'ion chlorure a été observé pour les deux formes cristallines de la EhCRT et pour la TcCRT. Il pourrait jouer un rôle dans la fonction lectine en orientant correctement le résidu histidine pour que la liaison hydrogène avec l'aspartate voisin soit optimale, participant ainsi au réseau de liaisons hydrogène du domaine lectine (**Figure 3.III.12 A et B**).

La présence de l'ion chlorure a été scrupuleusement vérifiée en enregistrant un signal anomal dans les deux formes cristallines de la EhCRT (**Figure 3.III.12 C**). Même si la présence d'un ion chlorure n'a pas été relevée dans les structures de Mo et HuCRT, ces deux protéines possèdent une histidine et une valine aux mêmes positions.

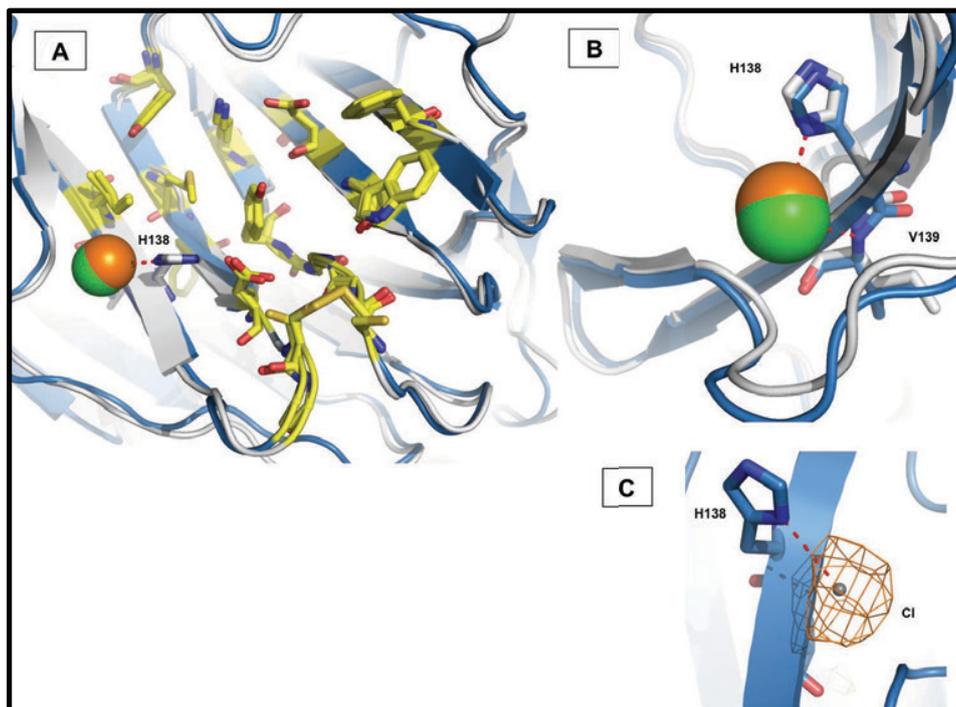


FIGURE 3.III.12 : SITE DE FIXATION DU CHLORURE DANS LA EHCRT ET LA TCRT

(A) Un ion chlorure est détecté au voisinage du site lectine (résidus conservés en jaune) de la EhCRT et de la TcCRT (respectivement en vert et orange). Il semble participer au réseau de liaisons hydrogène dans le site lectine en orientant correctement l'histidine 138 pour optimiser une liaison avec l'aspartate à proximité. (B) Un grossissement du site de fixation du chlore montre qu'il est coordonné par la chaîne latérale de l'histidine 138 et la chaîne principale de la V139 (EhCRT) ou de l'I139 (TcCRT). (C) La présence de l'ion chlorure a été confirmée en enregistrant un signal anormal une longueur d'onde de 2 Å (longueur d'onde la plus proche du seuil d'absorption du chlore (4,3929 Å) permise par la ligne de lumière) pour les deux formes cristallines de la EhCRT. La carte de différence anormale calculée (orange) est ici contourée à 5σ .

III.3. Etude de la dynamique du domaine P et de l'hélice C-terminale

Après avoir identifié les régions strictement conservées des calréticulines nous apportant des informations sur la fonction chaperon, nous nous sommes intéressés à l'étude des régions flexibles pour éclaircir les différents mécanismes impliqués dans la stabilisation ou le mouvement de ces régions et les relier potentiellement à leurs fonctions.

III.3.a. Flexibilité de l'extrémité de l'hélice C-terminale

Nous avons été en mesure de caractériser la flexibilité de l'hélice C-terminale, en observant la déviation de la chaîne principale de l'hélice dans les différents monomères de calréticuline. En effet, de légères fluctuations associées à une augmentation de l'agitation thermique sont observées à l'extrémité de l'hélice C-terminale des calréticulines et débutent lorsque l'hélice

se détache physiquement du domaine globulaire (**Figure 3.III.13**). Cette flexibilité doit être amplifiée sur la totalité de l'extrémité C-terminale.

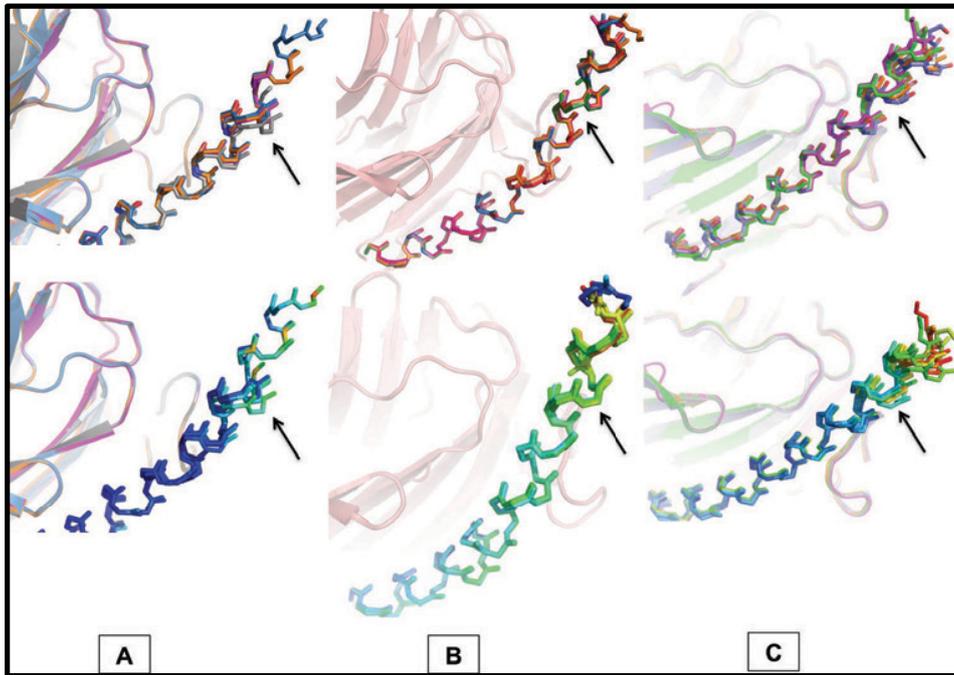


FIGURE 3.III.13 : FLEXIBILITE DE L'EXTREMITE C-TERMINALE DES CALRETICULINES

La superposition des différents monomères de calréticulines contenus dans l'unité asymétrique pour la HuCRT (A), la TcCRT (B), la EhCRT (C) montre une divergence de la chaîne principale de l'hélice plus ou moins importante selon les espèces. Alors que la position de l'hélice est strictement conservée au contact du domaine globulaire, elle commence à diverger au niveau de la position indiquée par la flèche noire. Cette divergence de la chaîne principale est corrélée à une augmentation des facteurs B d'agitation thermique (figures du bas).

Nous avons analysé plus finement les interactions impliquées dans la stabilisation de l'hélice au niveau de cette région se dessinant comme une charnière. L'hélice semble principalement stabilisée à travers des liaisons non covalentes impliquant le glutamate 354 et dans le domaine globulaire, la lysine 98 et la thréonine 173 (sérine pour EhCRT). Dans la EhCRT et la HuCRT une arginine et une lysine (358) respectivement maintiennent l'hélice par une liaison hydrogène impliquant l'aspartate 100. Enfin à la base de l'hélice une interaction hydrophobe importante qui avait été une des cibles de notre stratégie de dissection, participe à l'interaction, elle implique la leucine L171 du domaine globulaire conservée chez toutes les espèces et le résidu 347 (un W pour les Mo et HuCRT, une Y pour la TcCRT et une L pour la EhCRT). Enfin les résidus F et W55 de la TcCRT et de la EhCRT, respectivement, stabilisent

l'interaction à travers des interactions hydrophobes avec la chaîne latérale aliphatique de la lysine 358. Au delà de ces interactions la position l'hélice n'est plus contrainte (**Figure 3.III.14**).

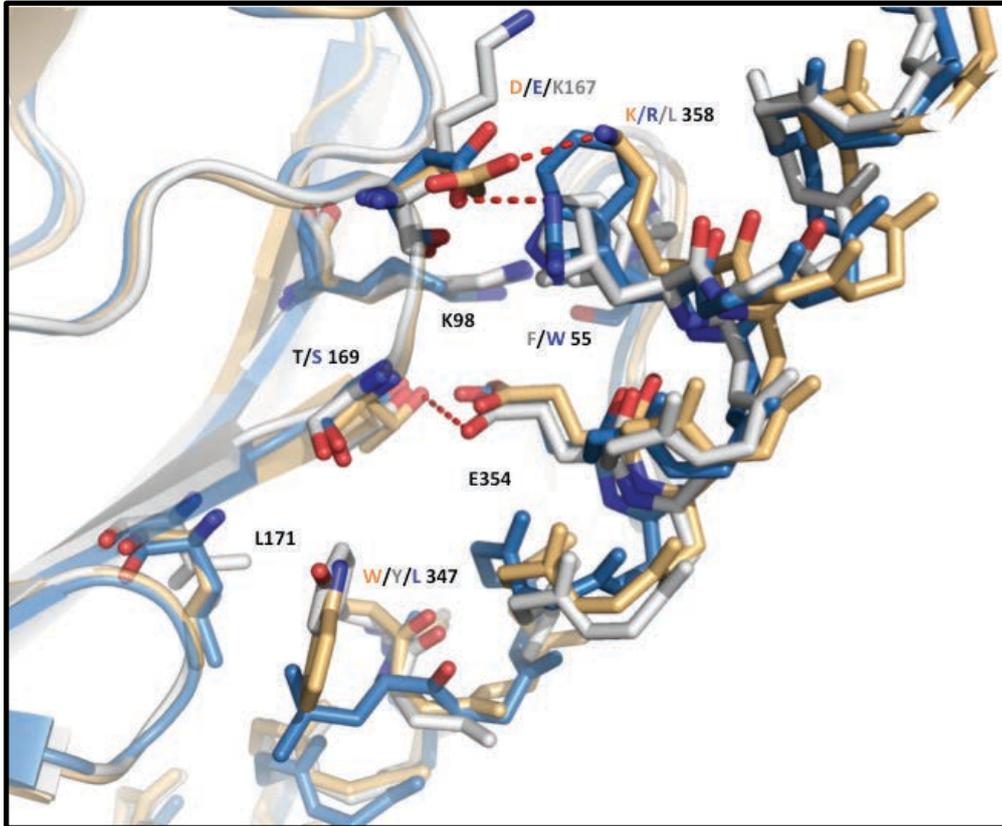


FIGURE 3.III.14 : STABILISATION DE L'HELICE C-TERMINALE DES CALRETICULINES PAR LE DOMAINE GLOBULAIRE

L'hélice C-terminale des calréticulines humaine (**orange**), de *E. histolytica* (**bleu**), de *T. cruzi* (**gris**) est stabilisée par des interactions non covalentes avec le domaine globulaire impliquant les chaînes latérales des résidus K/R358 (Eh/HuCRT), E354 dans l'hélice et T/S169, K98, E167 (K dans la TcCRT) dans le domaine globulaire. En plus, un résidu hydrophobe conservé dans le domaine globulaire (L171) est protégé par des interactions de van der Waals avec une chaîne latérale hydrophobe (W/Y/L 347). De même les résidus aromatiques F/W56 (Hu-TcCRT/EhCRT) sont protégés du solvant par les chaînes latérales K/R/L358 (Hu/Eh/TcCRT). La numérotation de référence est celle de HuCRT.

Les observations dans le cristal nous ont permis de pointer une légère flexibilité de l'hélice C-terminale définissant une charnière potentielle. Afin d'étudier plus précisément ce phénomène dans un environnement non contraint par des interactions protéiques, nous pouvons analyser les mesures de diffusion des rayons X aux petits angles sur différents fragments de la TcCRT. Ces mesures ont été effectuées sur la ligne BM29 de l'ESRF sur les fragments suivants : le fragment cristallisé TcCRT Δ Pter Δ 288, la protéine complète TcCRTFull et celle possédant un domaine P de la même longueur que dans la structure de la CRT de souris, TcCRT Δ P, et enfin la TcCRT Δ P Δ 281, possédant le même domaine P que la TcCRT Δ P mais une hélice C-terminale plus courte (**Figure 3.III.16**).

Les courbes de diffusion mesurées à 4 concentrations pour chaque fragment étaient de bonne qualité, le domaine de Guinier ne montrait pas de phénomène d'agrégation, bien qu'une diminution de l'intensité de diffusion dans cette région à haute concentration suggérait des interactions répulsives entre les particules. De même, le rayon de giration déterminé par l'approximation de Guinier n'est pas dépendant de la concentration. Pour chacun des fragments, nous avons créé une courbe de diffusion « optimisée » en utilisant le domaine de Guinier (0-1 nm⁻¹) de la courbe enregistrée à la plus basse résolution, pour minimiser les interactions inter-particulaires, le domaine de Porod (1-2,5 nm⁻¹) et la région des grands angles (>3 nm⁻¹) de la plus haute concentration pour avoir un meilleur rapport signal sur bruit. Un exemple de courbe de diffusion utilisée pour la suite du traitement des données est présenté sur la **figure 3.III.15**

En suivant la méthode décrite dans la section Matériels et Méthodes, les courbes de diffusion ont été utilisées pour calculer les courbes de Guinier et les fonctions de distribution de distance pour chaque fragment (**Figure 3.III.16**). Ces opérations nous ont permis, en plus d'estimer la qualité des données, de calculer un rayon de giration (Rg) et une distance maximale (Dmax). Les valeurs obtenues et présentées sur la **figure 3.III.16** sont cohérentes avec les différents fragments. En effet, le Rg décroît de 3,75 à 2,1 nm en passant de la TcCRTFull au fragment cristallisé, tandis que la Dmax diminue de 13,1 nm à 7,1 nm.

Parallèlement, les fonctions de distribution de distance nous renseignent sur la forme de l'objet en solution. Ici aussi, les résultats sont cohérents avec nos attentes. La courbe de distribution des distances du fragment cristallisé est une gaussienne quasi parfaite, caractéristique d'une protéine globulaire, alors que pour les fragments possédant un domaine

P de longueur moyenne, la courbe présente un épaulement associé à la présence d'un second domaine allongé. Enfin, pour le fragment complet, nous observons une courbe ayant une apparence se rapprochant de ce que l'on pourrait observer pour une protéine possédant 2 domaines globulaires, à savoir une double gaussienne (**Figure 3.III.16**).

Ces paramètres biophysiques nous ont permis, en utilisant le programme GASBOR, de modéliser une enveloppe correspondant à la forme globale de l'objet en solution. Ces enveloppes sont à utiliser avec prudence puisqu'en effet une expérience de SAXS n'apporte aucune information tridimensionnelle et elles ne résultent donc que de modélisation, notamment pour des protéines flexibles pour lesquelles le logiciel converge avec difficulté, comme le montre l'enveloppe très large proposée pour la protéine entière. Néanmoins, les enveloppes permettent de se faire une idée de la forme de l'objet étudié.

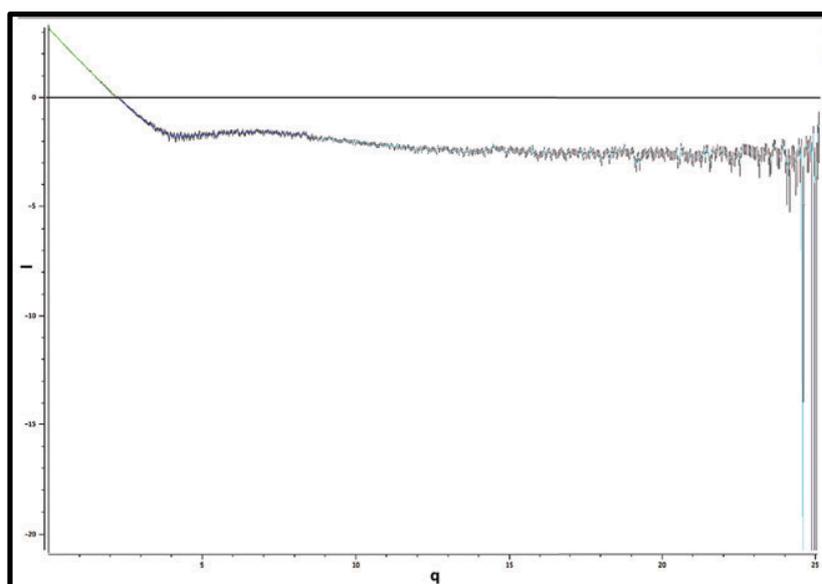


FIGURE 3.III.15 : COURBE DE DIFFUSION OPTIMISEE DU FRAGMENT TcCRTΔPterΔ288

La courbe de diffusion utilisée pour traiter les données SAXS du fragment TcCRTΔPterΔ288 a été générée à partir du domaine de Guinier de la courbe de diffusion à basse concentration (en vert), le domaine de Porod de la courbe à concentration intermédiaire (bleu foncée) et la région des grands angles de la courbe à haute concentration (bleu ciel)

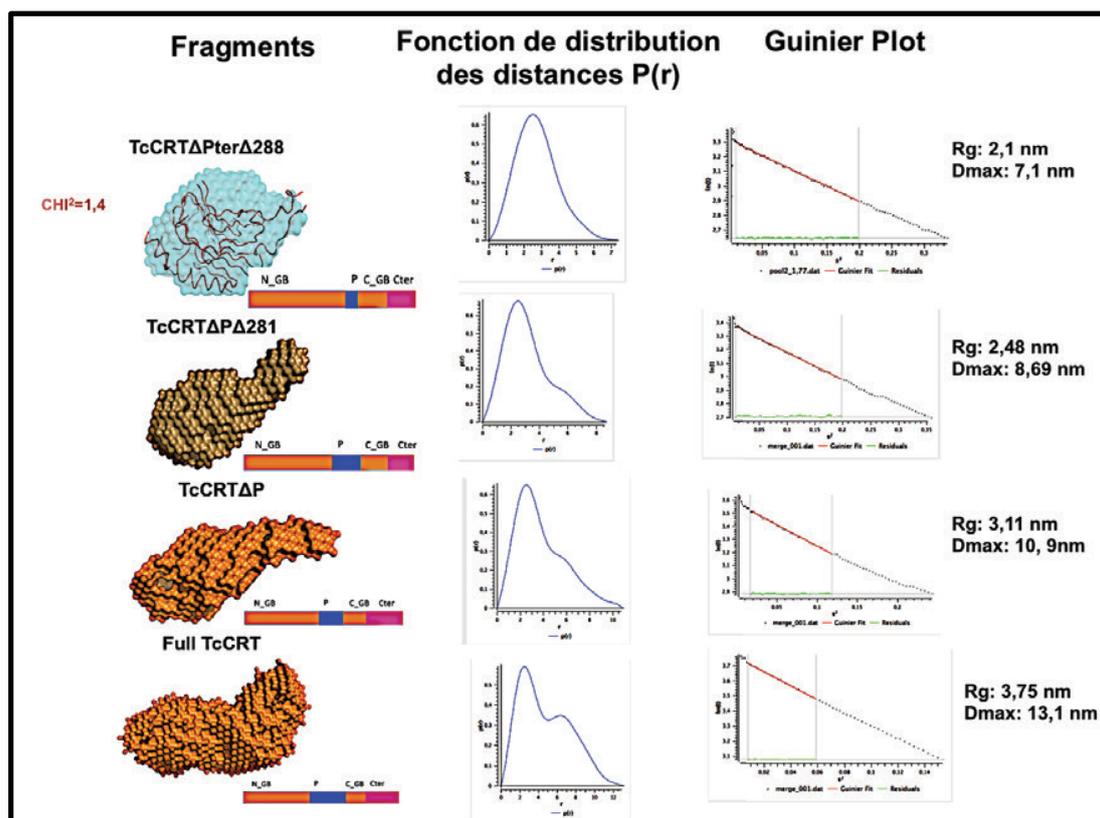


FIGURE 3.III.16 : TRAITEMENT DES DONNEES SAXS SUR LES FRAGMENTS DE CALRETICULINES

La diffusion des rayons X enregistrée aux petits angles sur les fragments de TcCRT (TcCRTFull, ΔP , $\Delta P\Delta 281$, $\Delta Pter\Delta 288$) a permis de déterminer des paramètres biophysiques cohérents avec chacun des fragments (Rg, Dmax, forme, suppression d'un domaine). Ces paramètres ont été utilisés dans des expériences de modélisation pour calculer des enveloppes correspondant à la forme globale approximative de la protéine en solution.

Bien que limité pour obtenir une information 3D, le SAXS est cependant un outil très puissant pour comparer une courbe de diffusion expérimentale et une courbe de diffusion calculée, en utilisant le logiciel Crysol. C'est cet outil que nous avons exploité pour étudier dans un premier temps la flexibilité du fragment cristallisé, TcCRT $\Delta P\Delta 288$, cas de figure le plus simple, car possédant le moins de régions flexibles, mais aussi une structure cristallographique comme référence.

Même si nous avons supprimé de nombreuses régions supposées flexibles dans ce fragment, la différence entre la courbe de diffusion en solution et celle calculée à partir de la structure suggère une différence locale de conformation en solution et une flexibilité résiduelle (**Figure 3.III.17**), confirmée par l'allure du Kratky plot, que j'ai choisi de ne pas présenter. Pour identifier la région à l'origine de cette divergence, nous avons utilisé le logiciel BUNCH de la

suite ATSAS. Ce logiciel permet de déplacer des protéines ou des fragments de protéines en blocs rigides pour faire correspondre une courbe de diffusion calculée à celle enregistrée. Par exemple, il permet de déplacer la structure de deux monomères pour obtenir un modèle correspondant à la courbe de diffusion d'un dimère en solution.

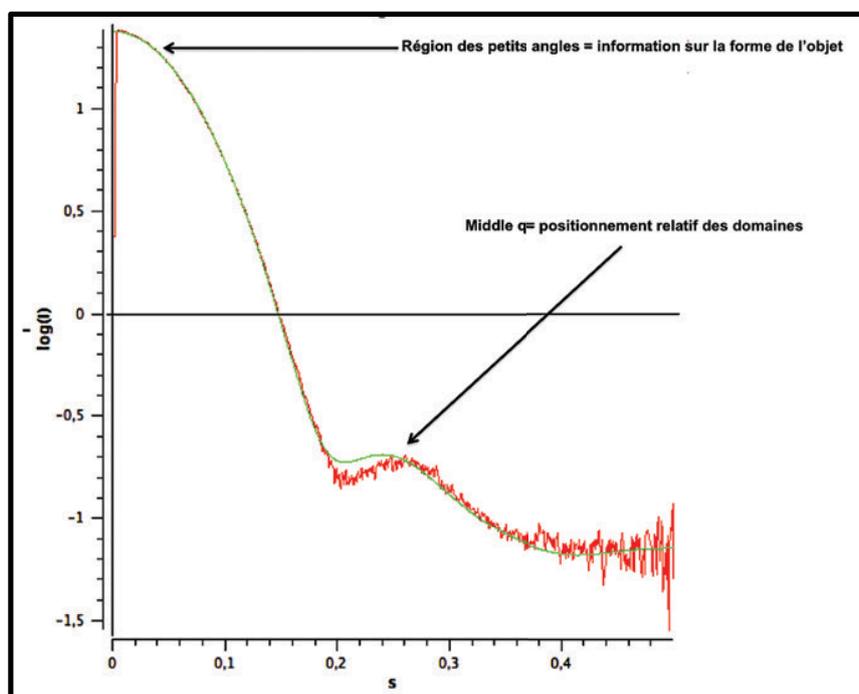


FIGURE 3.III.17 : COMPARAISON DES COURBES DE DIFFUSION CALCULEE ET MESUREE POUR LE FRAGMENT TCRTΔPTERD288

La comparaison, en utilisant le logiciel CRY SOL, de la courbe de diffusion du fragment en solution et la courbe calculée à partir de la structure montre une différence dans le domaine de Porod, suggérant une position différente d'un domaine. La superposition parfaite de la courbe dans le domaine de Guinier ne suggère aucune modification majeure de la conformation de la protéine en solution.

J'ai utilisé le logiciel BUNCH pour tester différentes charnières autour desquelles l'hélice C-terminale et le domaine P pouvaient s'articuler, et j'ai recommencé plusieurs fois l'opération pour observer si nous pouvions converger vers une solution unique. Il se trouve que nous n'avons jamais été en mesure d'obtenir une solution unique, même si de nombreuses solutions présentées figure 3.III.18 permettaient d'obtenir une très bonne corrélation des deux courbes. Ces solutions correspondaient à des positions différentes de l'extrémité de l'hélice C-terminale après une charnière située en amont de celle que nous avons observée (au niveau du résidu K351), confirmant la mécanique suggérée après l'observation des structures. Aucun mouvement de la fraction du domaine P restant ne permet d'ajuster les courbes expérimentale et calculée.

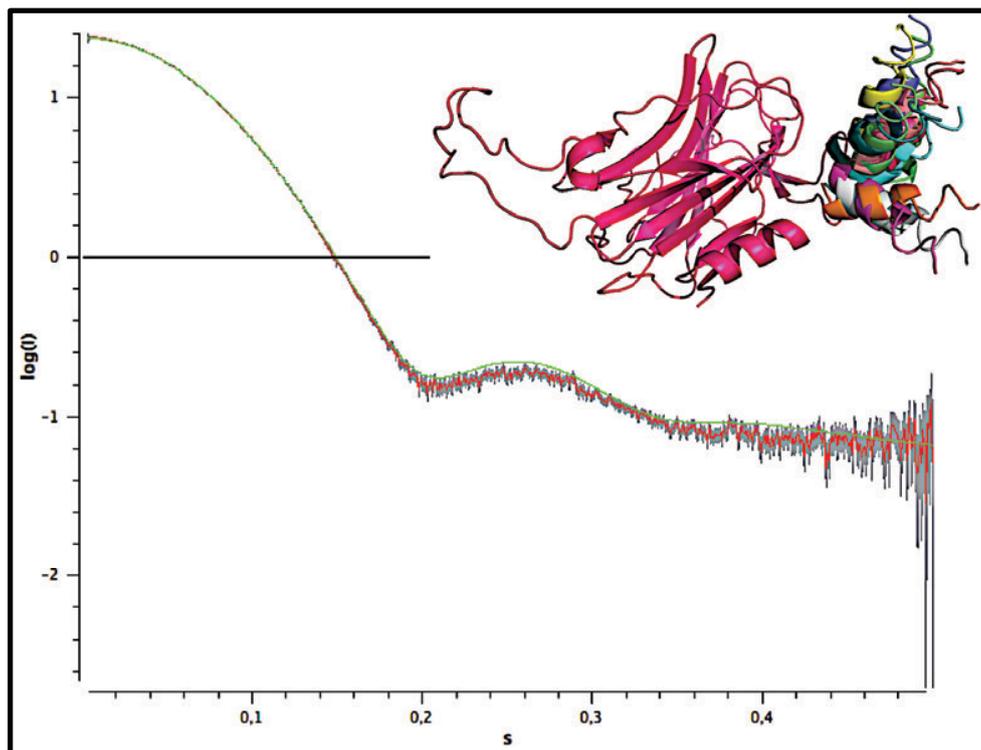


FIGURE 3.III.18: COMPARAISON DES COURBES DE DIFFUSION CALCULEES A PARTIR DE MODELES ISSUS DE BUNCH ET CELLE MESUREE POUR LE FRAGMENT TcCRTΔPterD288

Les différentes positions proposées par le logiciel BUNCH pour l'extrémité de l'hélice C-terminale permettent de faire correspondre les courbes de diffusion mesurée et calculée du fragment TcCRTΔPterD288, notamment au niveau du domaine de Prod. Dans ces modèles, l'hélice devient mobile après le résidu K351.

III.3.b. Flexibilité du domaine P

Comprendre la dynamique de mouvement et du repliement du domaine P était également intéressante pour déterminer les mécanismes à travers lesquels il peut s'associer à ERp57 par exemple ou aux substrats non glycosylés et déterminer la conformation que ce domaine adopte en solution. En effet, le domaine P est souvent représenté ou modélisé en position relativement fermée en solution. En utilisant le logiciel CRY SOL comme précédemment, nous avons tout d'abord pu montrer que cette représentation n'était pas du tout cohérente avec les mesures en solution. En effet, les courbes de diffusion mesurée et calculée ne correspondent absolument pas dès le domaine de Guinier, signifiant que la protéine dans cette conformation présente un R_g inférieur aux mesures. Pour déterminer la ou les conformations principales du domaine P en solution, nous avons utilisé une stratégie équivalente à celle appliquée à l'hélice C-terminale. Nous avons imaginé des charnières probables autour desquelles le domaine P pouvait s'articuler, et nous avons utilisé le logiciel BUNCH pour déplacer les fragments de structure en blocs rigides autour de ces charnières (**Figure 3.III.19**).

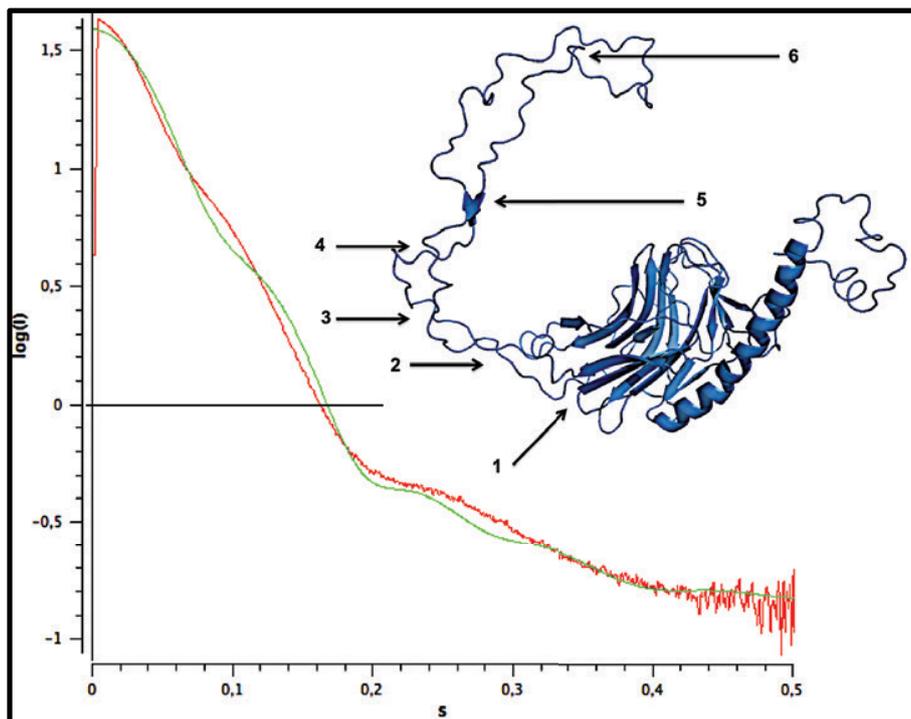


FIGURE 3.III.19 : VALIDATION D'UN MODELE DE LA TcCRTFULL

La modélisation de la TcCRTFull, comme par exemple ici en utilisant le serveur Phyre², conduit systématiquement à la représentation du domaine P en position fermée au dessus du domaine lectine. La courbe de diffusion calculée à partir de ce modèle ne correspond pas aux mesures et montre que notre ce modèle est trop compact. Pour étudier la flexibilité du domaine P, nous avons défini les 6 charnières potentielles indiquées ici.

Pour étudier la flexibilité de la TcCRTFull, nous avons imaginé un modèle hybride composé de la structure obtenue en cristallographie aux rayons X et des fragments du domaine P issus de la modélisation. Pour cette étude, nous nous sommes focalisés sur les domaines de Guinier et de Porod (0 à 3nm^{-1}). En effet la protéine étant flexible, il est difficile de faire converger les données à haute résolution, en absence de conformation majoritaire en solution. Nous nous sommes également vite rendu compte que la conformation du domaine P ne suffisait pas à expliquer les données expérimentales et que les modèles les plus proches de la réalité associaient des conformations différentes du domaine P et de l'hélice C-terminale.

Plusieurs modèles nous ont permis d'expliquer la courbe de diffusion enregistrée. Tous suggèrent une conformation plus ouverte du domaine P, dont les mouvements se font essentiellement en un seul bloc autour de la charnière 2. Ce domaine semble donc être animé d'un mouvement de levier et non s'enrouler sur lui même (**Figure 3.III.20**).

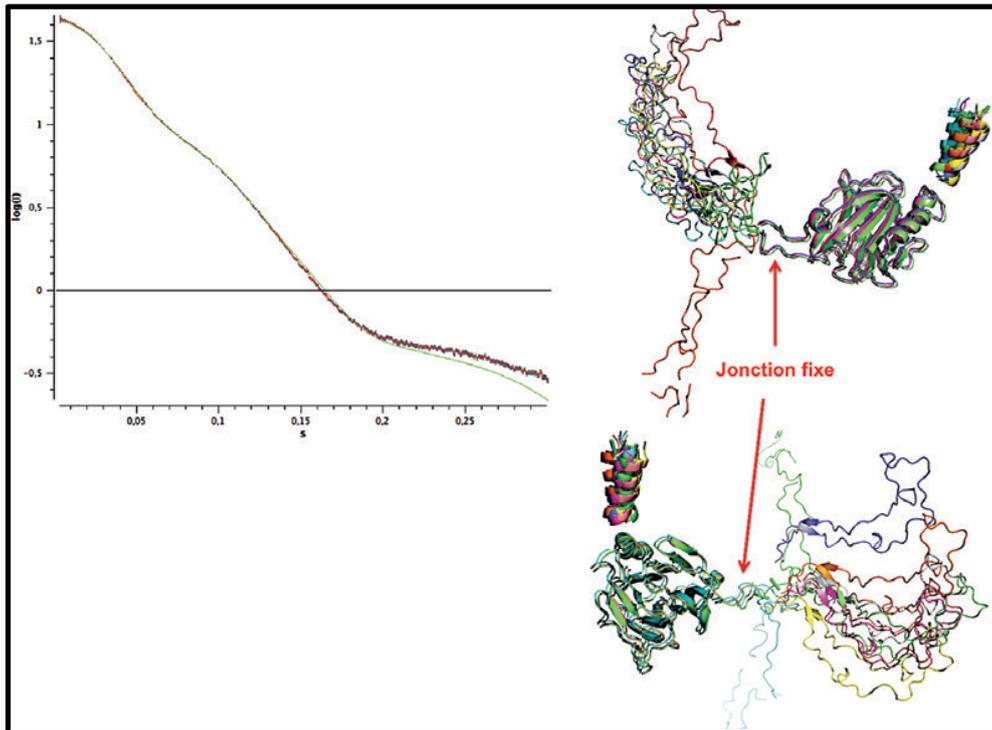


FIGURE 3.III.20 : CONFORMATIONS DU DOMAINE P EN SOLUTION

Les modèles obtenus avec le logiciel bunch en utilisant les courbes de diffusion enregistrées aux petits angles sur les fragments TcCRTFull, ΔP , $\Delta P\Delta 281$ et $\Delta P_{ter}\Delta 288$ montrent une conformation ouverte du domaine P en solution. Le domaine P semble s'articuler principalement autour de la charnière 2. Avant cette charnière, la jonction du domaine P semble rester stable.

III.3.c. Stabilité de la jonction du domaine P

Comme je viens de le présenter, le domaine P semble s'articuler autour d'une charnière située juste après la jonction reliant le domaine globulaire au domaine P (notée j sur la **figure 3.III.3**). Nous nous sommes alors demandé si effectivement cette jonction était conservée et demeurerait stable dans les différentes structures de calréticulines. La jonction inclut deux portions de 4 acides aminés. En comparant les différentes structures, nous pouvons maintenant expliquer comment cette petite portion module l'orientation relative du domaine P. La structure de la jonction est extrêmement conservée entre les TcCRT, MoCRT et HuCRT. Sa structure, plus variable dans la EhCRT résulte potentiellement de la présence d'un résidu alanine créant une charnière flexible (**Figure 3.III.21**).

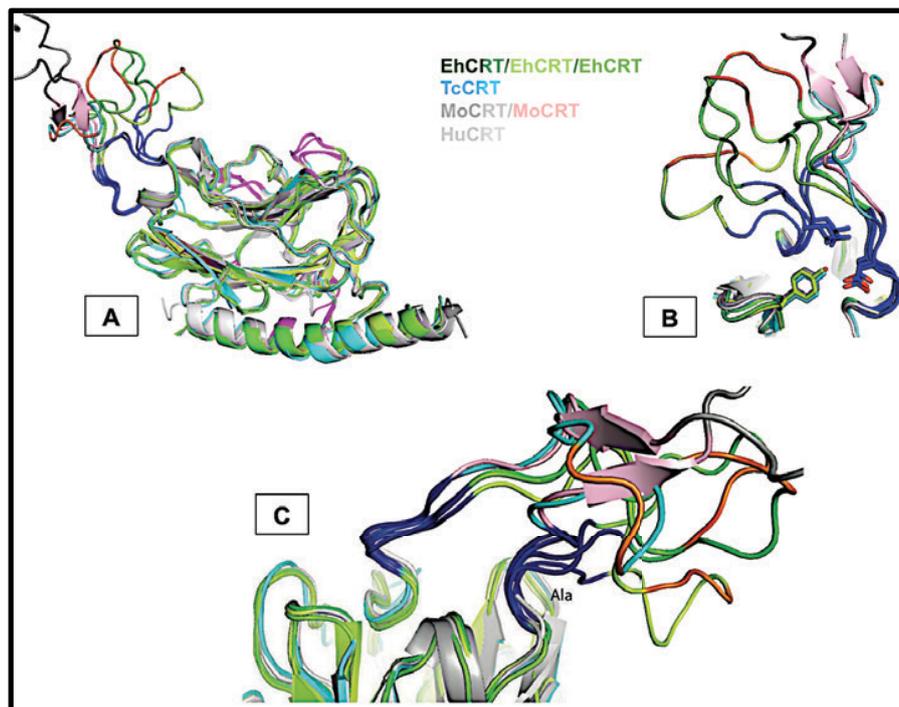


FIGURE 3.III.21: ORIENTATION DE LA JONCTION ENTRE LE DOMAINE P ET LE DOMAINE GLOBULAIRE DANS LES CALRETICULINES

La conformation de la jonction entre le domaine globulaire et le domaine P est hautement conservée dans les différents monomères des calréticulines observés dans les cristaux (**A** et **B**). Nous observons une légère divergence pour deux monomères de la EhCRT, pouvant s'expliquer par la présence d'une alanine (à la place d'une leucine ou proline) apportant davantage de flexibilité à la jonction (**C**).

Dans les calréticulines, la partie N-terminale de la jonction est très hydrophobe. Le premier résidu (F/L/M) ancre la jonction perpendiculairement au domaine globulaire (**Figure 3.III.22 (C)**). Le résidu hydrophobe suivant (I/L) contribue au cœur hydrophobe de la jonction (196, 201, 203). Les deux derniers résidus sont souvent deux prolines, à l'exception de la EhCRT et TcCRT où une seule proline est conservée. Dans le cas de la EhCRT, une alanine est présente à la place de la première proline conservée, contribuant à l'augmentation de la flexibilité comme je l'ai déjà mentionné. La dernière proline introduit un coude qui amplifie le mouvement de charnière de l'alanine (**Figure 3.III.22 (A) et 3.III.20**).

L'extrémité C-terminale de la jonction démarre par un premier acide aminé variable (P/Y/E). L'acide glutamique, présent dans la TcCRT, interagit avec le premier résidu du domaine P, le résidu basique R204 (**Figure 3.III.22 (A)**). Un résidu acide conservé interagit systématiquement avec le groupement hydroxyle de la chaîne latérale de la tyrosine conservée du domaine globulaire. Ce résidu conservé est également impliqué dans une interaction avec

la chaîne latérale de la lysine présente dans le domaine globulaire (K151 dans Mo et HuCRT, K144 dans EhCRT), à l'exception de la TcCRT (**Figure 3.III.22**). Dans le cas de la TcCRT et EhCRT, le dernier résidu acide de la partie C-terminale de la jonction interagit avec des résidus basiques du domaine globulaire : K88 (TcCRT) et K79 (EhCRT) (**Figure 3.III.22** (C)).

Le premier résidu basique du domaine P de la CRT est attiré par un cluster de charges négatives présentes dans la petite hélice du domaine globulaire précédant le domaine P (mise à part dans la EhCRT car cette région est déplacée et la distance drastiquement augmentée) (**Figure 3.III.22** (C)).

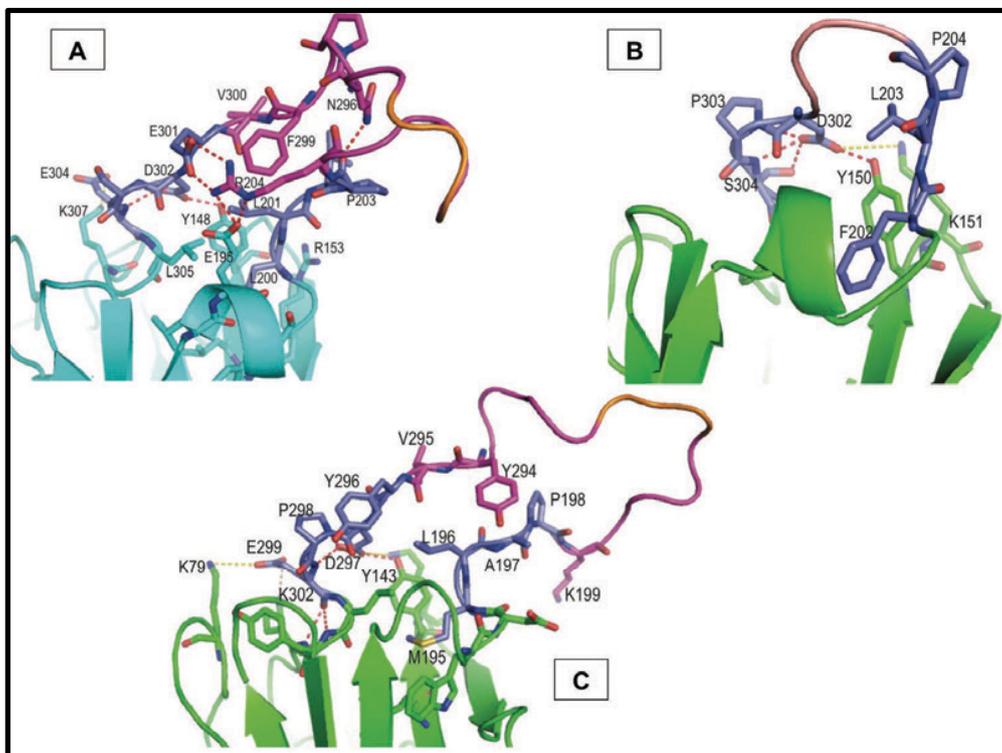


FIGURE 3.III.22: INTERACTIONS STABILISANT LA JONCTION DU DOMAINE P DANS LES CRTS
Représentation des liaisons impliquées dans le maintien de la jonction du domaine P dans la TcCRT (A), HuCRT (B) et EhCRT (C)

Dans le cas particulier de la EhCRT, nous avons pu observer différentes conformations du domaine P résiduel pour les monomères présents dans l'unité asymétrique. Dans l'un d'entre eux, ce domaine passe d'une position ouverte à fermée. L'étude de la réorganisation de la jonction dans ce cas particulier améliore notre compréhension des mécanismes moléculaires mis en jeu lorsque le domaine P se referme sur un substrat. Ici, le résidu M195 décrit précédemment n'est plus ancré dans le domaine globulaire et la position de son C α a bougé

de 5,2 Å, tandis que d'autres déplacements plus importants sont observés. Par exemple les isoleucines 201 et 289 se déplacent de 18 et 22 Å respectivement et 26 Å pour la sérine présente dans le linker.

Ce changement de conformation complexe ne s'explique pas uniquement par un mouvement charnière. En plus des déplacements, il se produit des réarrangements inhabituels, comme par exemple de la Y143 à C156 où la chaîne polypeptidique vrille au niveau du résidu G145. Cela induit la réorientation de chaînes latérales à 180° similaire à une insertion/délétion ponctuelle sur un brin. L'effet le plus sévère se produit de K146 à E153.

Ce changement de conformation inhabituel est associé au déroulement de l'hélice 1 et au déplacement de l'extrémité N-terminale de la jonction. Lors de ce réarrangement, la distance entre les résidus 175 et 178 augmente de 5,4 à 9 Å. La jonction N-terminale commence maintenant au niveau d'un petit feuillet β dans une nouvelle conformation qui est maintenue par des liaisons hydrogène entre les chaînes principales des résidus L196 et N149 et la chaîne principale de l'alanine 197 avec les chaînes principale et latérale du résidu E200 d'un côté et de l'autre côté, la jonction est stabilisée par une liaison hydrogène entre les chaînes principales des résidus M195 et N209. Enfin une interaction hydrophobe entre les résidus L196 et I150 participe à la stabilisation de la jonction (**Figure 3.III.23 (C)**).

Un autre effet charnière se produit au niveau de l'extrémité C-terminale de la jonction au niveau du résidu E299. Ce résidu est attiré par des interactions électrostatiques avec les résidus K79 et K302 du domaine globulaire. Des liaisons hydrogène entre la chaîne latérale de la N173 et les groupements hydroxyle et carbonyle de Y296 et P298, respectivement, sont également maintenus pendant le changement de conformation, bien que la position de Y296 se décale de 4,5 Å.

La cystéine libre, comme le pont disulfure, restent stables à l'extrémité du domaine lectine et semblent jouer un rôle crucial dans sa stabilisation. Seulement deux glycines (G132-133) au voisinage du pont disulfure se déplacent dans la nouvelle conformation (**Figure 3.III.23 (B)**).

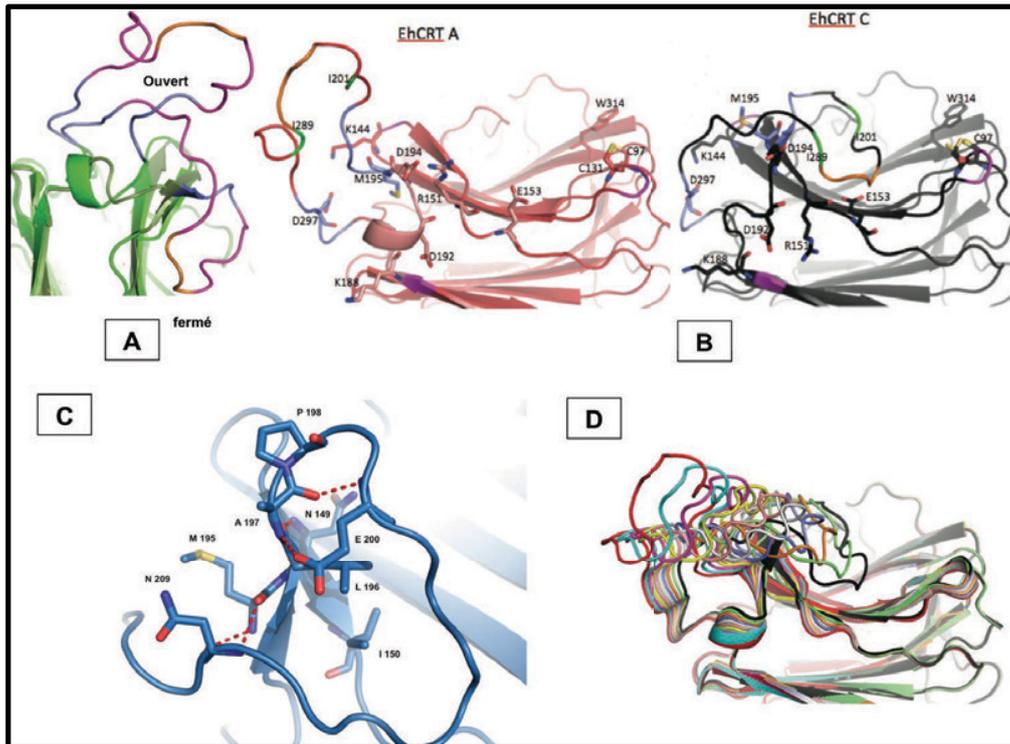


FIGURE 3.III.23: DESCRIPTION MOLECULAIRE DE LA FERMETURE DU DOMAINE P

(A) L'observation des chaînes A et C dans l'unité asymétrique des cristaux de EhCRT montre la transition du domaine P d'une position ouverte à une position fermée. (B) Ce réarrangement conformationnel induit un déplacement de la chaîne principale et une réorganisation des chaînes latérales des résidus qui sont impliqués dans de nouvelles interactions. (C) Un nouveau réseau de liaisons hydrogène participe à la stabilisation de la jonction de la chaîne C de la EhCRT dans une configuration où le domaine P est fermé. (D) Une interpolation entre les conformations ouverte et fermée permet de mieux visualiser les changements conformationnels assurant la transition entre ces deux états.

III.4. Stabilisation du domaine P dans le cristal : mise en évidence d'un site d'interaction des peptides indépendants du glycanne ?

Mises à part les interactions visibles avec le glucose au niveau du domaine lectine, les études structurales menées ont fourni des exemples d'interaction du domaine lectine avec des molécules voisines dans le cristal. Cependant, ces contacts se produisent exclusivement avec des segments flexibles des molécules voisines et exposant des résidus hydrophobes. Par conséquent, ces contacts pourraient être interprétés comme similaires à des interactions de type chaperonne.

Le premier exemple de ces interactions est illustré dans l'empilement cristallin de la TcCRT où nous sommes dans le cas d'une interaction hybride avec le glucose. L'interaction avec le

peptide est alors limitée au bord du domaine lectine et implique les résidus F73, C105, C136, S138, W320 et V322. Du côté du « peptide », cette interaction implique les résidus hydrophobes I206 et 294 (du P), les glycines du linker, les P295 et 297 (de la jonction) et V207. Quelques interactions polaires additionnelles se produisent sur le bord du domaine avec les résidus D298 et S138 (**Figure 3.III.24**).

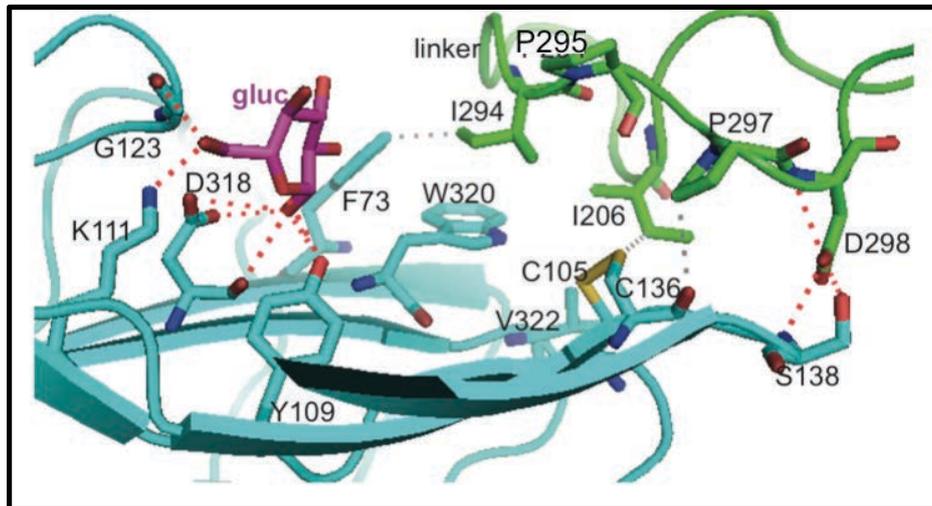


FIGURE 3.III.24 : DESCRIPTION DE L'INTERACTION ENTRE LE DOMAINE P ET LE DOMAINE LECTINE DANS LA TcCRT

Le second exemple provient des cristaux de EhCRT. Dans ce cas, en absence de glucose, la surface de contact est plus grande et implique des résidus du site de fixation du glucose en plus des résidus du bord du domaine lectine. Dans ce cas de figure, le résidu D129 stabilise la chaîne principale de deux résidus hydrophobes exposés. Le groupement hydroxyle de la Y101 stabilise également le carbonyle de la chaîne principale de la G204 présente dans le linker. A l'exception de ces deux interactions, les autres sont toutes de type van der Waals. Du côté du site lectine, les résidus S134, I140, L149, le pont disulfure (C97 et 131) et W314 participent à des interactions stabilisant les chaînes latérales des résidus I201 et 289, A290, P292, Y294 et le linker côté peptide (**Figure 3.III.25**).

Nous pouvons noter, de manière intéressante, que le pont disulfure et la chaîne latérale du tryptophane sur le bord du site lectine sont toujours impliqués dans des contacts hydrophobes (avec I289 dans la CRT, P171 dans les chaînes A et C de la EhCRT et P292 dans la chaîne B). Ces observations sont corrélées à celles précédemment décrites avec le domaine mobile de type peptide situé à l'extrémité N-terminale des CRTs de mammifères, même si aucune chaîne latérale hydrophobe n'est impliquée dans ce dernier cas (**Chouquet et al., 2011**).

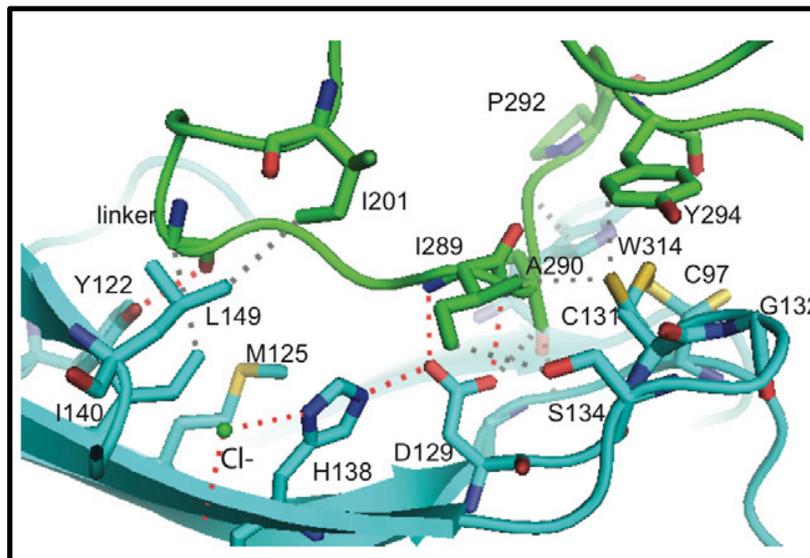


FIGURE 3.III.25: DESCRIPTION DES INTERACTIONS PERMETTANT LA STABILISATION DU DOMAINE P PAR LE DOMAINE LECTINE DANS LA EHCRT

Enfin, une seule exception a pu être observée, celle impliquant le site lectine de la chaîne A et le monomère C de la EhCRT, pour lequel le domaine P se referme. Cette interaction implique une surface d'interaction étendue avec le site lectine et le domaine P tronqué. Une molécule de glucose est piégée entre les deux domaines globulaires (**Figure 3.III.26**). De nombreux contacts polaires et interactions électrostatiques participent à ce contact. Deux résidus basiques (R170 et K188) interagissent avec des résidus acides du domaine lectine (D129 et E119 respectivement), avec un pont salin entre K188 et E119. Les résidus basiques du domaine lectine sont également impliqués dans des interactions polaires : H138 avec E186, R136 avec E183, K146 avec E191 et K317 avec E332. E335 et K178 interagissent avec deux glycines consécutives (G132-133). Un cluster de résidus acides (190 à 192) est proche de la paire de lysines 146 et 199 du site lectine. Trois résidus hydrophobes sont exposés : I168, V176 et I331. Ces résidus sont assez proches du pont disulfure (entre 3,7 et 4,9 Å). Sur l'autre face, la proline 171 est impliquée dans des contacts de type van der Waals avec le W314. Enfin le résidu I289 est stabilisé par une interaction avec les atomes de la chaîne principale d'un acide glutamique E200, précédent I201 et D202. Une majorité de ces résidus de la chaîne C participant à l'interaction ne sont pas dans leur conformation ou leur orientation native et pourraient donc mimer une protéine dans une conformation altérée dans le contexte de la fonction chaperonne indépendante du glycan (**Figure 3.III.26**).

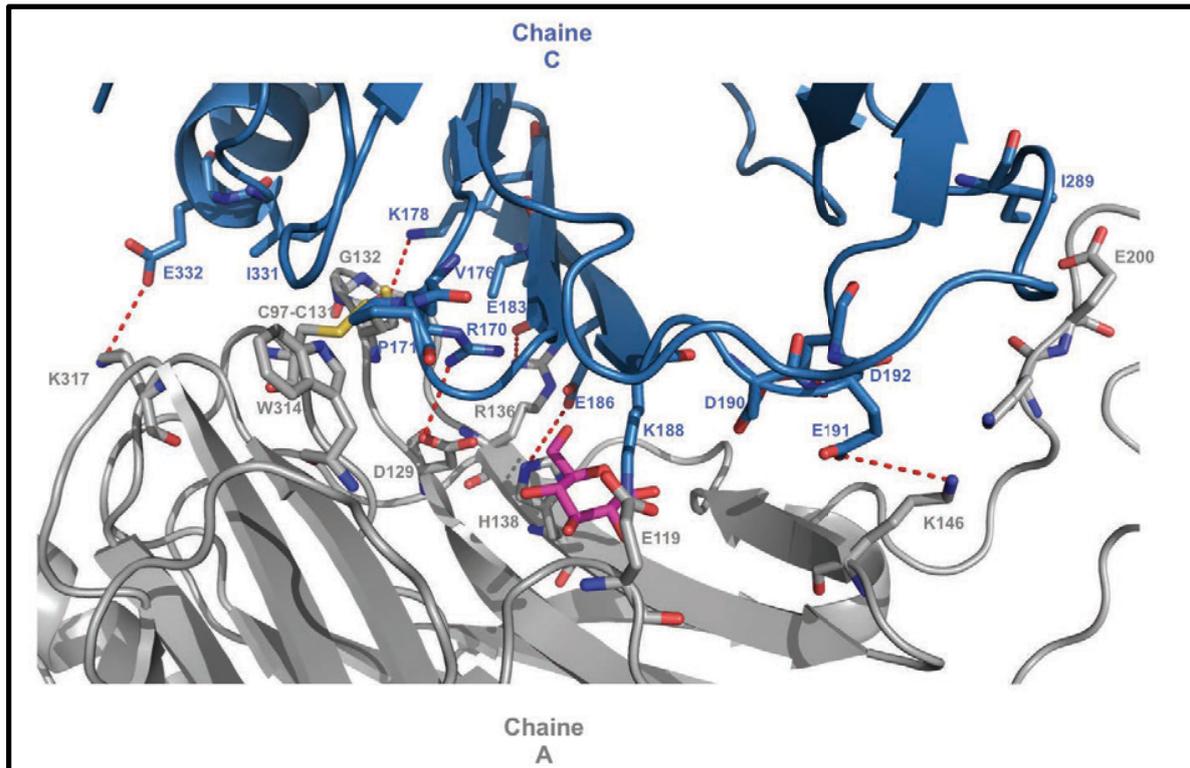


FIGURE 3.III.26: UNE EXCEPTION : INTERACTION ENTRE LE DOMAINE GLOBULAIRE DE LA CHAÎNE C ET LE DOMAINE LECTINE DE LA CHAÎNE A DE LA EHCRT DANS L'UNITÉ ASYMETRIQUE

Les résidus déjà décrits et impliqués dans cette interaction présentent tous une conformation ou une orientation non native dans la chaîne C. Nous pouvons observer une molécule de glucose (en rose) coincée entre les deux monomères.

Chapitre 2 : Les calréticulines comme récepteurs de C1q

IV. Caractérisation des interactions entre les fragments de CRT et C1q

IV.1. Identification des domaines impliqués dans l'interaction

L'interaction des calréticulines de parasite et humaine avec C1q a déjà été largement étudiée. Dans le cas de la calréticuline humaine, il a été montré par SPR que le domaine globulaire de la CRT pouvait interagir seul avec C1q, au contraire du domaine P (**Païdassi et al., 2011**). Cependant nos collaborateurs à Santiago ont affirmé que le domaine P était essentiel à l'interaction et l'infektivité dépendante de C1q (**Ferreira et al., 2004a**). Pour la EhCRT, aucune dissection de la protéine n'a été effectuée. Nous avons souhaité comparer les interactions de ces fragments avec C1q pour déterminer si effectivement dans le cas des calréticulines de parasite le domaine P (ou quelle région de ce domaine) était nécessaire. Les fonctions des calréticulines de ces deux parasites se rapprochent de celles (signal « eat-me » et récepteur montant à la surface de la cupule phagocytaire) portées simultanément par la protéine humaine. De ce fait identifier les régions impliquées dans l'interaction et partagées exclusivement entre les Tc et HuCRT ou entre Eh et HuCRT pouvait nous aider à comprendre les mécanismes de reconnaissance mis en jeu dans la phagocytose et l'élimination des cellules apoptotiques mais également créer l'espoir d'identifier des sites d'interactions spécifiques à chaque espèce, permettant de lutter contre les parasites sans affecter les interactions HuCRT-C1q.

Pour cela les fragments TcCRTFull, TcCRT Δ P, TcCRTCrys (= Δ Pter288), TcCRTNGB (sans domaine P), EhCRTFull, EhCRT Δ P, EhCRTNGB, HuCRTFull et HuCRTNGB ont été fixés sur un support solide comme décrit dans la section Matériels et Méthodes, et incubés avec C1q à une concentration de 10 nM. La révélation de l'interaction s'effectue grâce à des anticorps primaires dirigés contre C1q, puis des anticorps secondaires couplés à la HRP.

Les courbes d'interactions présentées **figure 3.IV.1** montrent que tous les fragments sont capables d'interagir avec C1q.

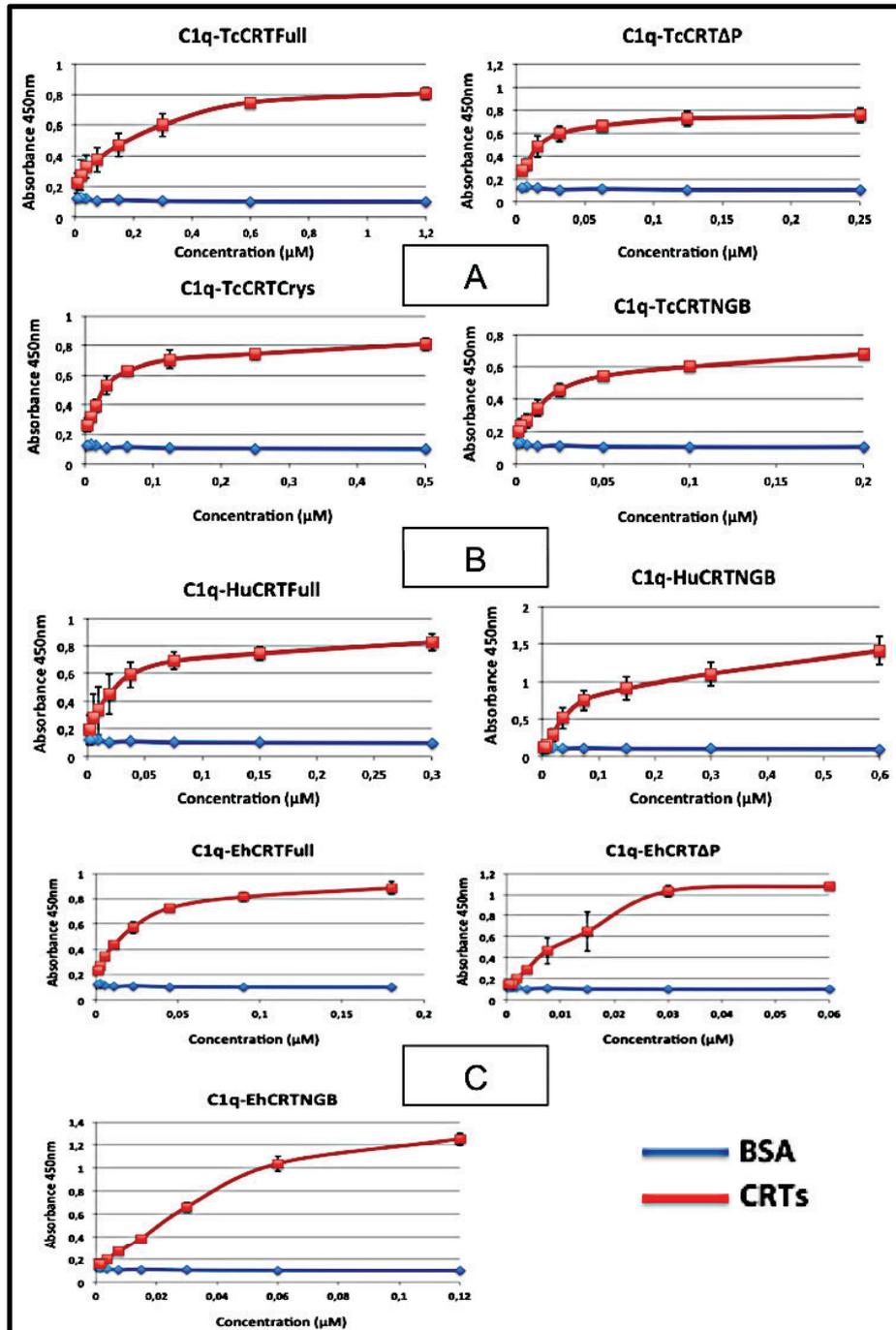


FIGURE 3.IV.1 : INTERACTIONS ENTRE C1Q ET LES FRAGMENTS DE CALRETICULINES (ELISA)

Les différents fragments de calréticulines (Tc, Hu, Eh en A, B et C, respectivement) ou la BSA ont été adsorbés sur une surface solide à huit concentrations différentes et incubés avec la protéine C1q à une concentration de 10 nM. L'interaction est révélée après lavage par l'ajout successif d'anticorps de lapin anti-C1q et d'anticorps dirigés contre les anticorps de lapin et couplés à la HRP. Pour que les résultats soient comparables, différentes plaques ont été réalisées en parallèle et la réaction de coloration a été stoppée au même temps. Deux expériences indépendantes ont été effectuées en triplicat, donc chaque point est la moyenne de 6 valeurs. Nous avons vérifié que les anticorps secondaires ne reconnaissent pas C1q.

De manière surprenante nous avons constaté que les fragments de EhCRT et de TcCRT ne possédant pas de domaine P semblaient interagir avec une meilleure affinité. En effet, plus le domaine P est court, plus le plateau est atteint à une faible concentration de calréticuline fixée. Il se pourrait que la flexibilité du domaine P gêne l'interaction. Auquel cas, nous n'aurions pas 100% des protéines adsorbées sur la surface capables d'interagir, contrairement aux fragments dépourvus de domaine P ou avec un domaine P raccourci (**Figure 3.IV.1 A et C**). Au contraire, pour la calréticuline humaine, nous avons dû doubler la concentration du domaine NGB pour atteindre le plateau. Cette observation pourrait suggérer une importance du domaine P dans l'interaction de la CRT humaine avec C1q au contraire des CRTs de parasites, même si cela avait été exclu par SPR (**Païdassi et al., 2011**) (**Figure 3.IV.1 B**).

Après avoir observé que le domaine globulaire des calréticulines seul était suffisant pour l'interaction avec C1q, nous avons voulu déterminer inversement quel domaine de C1q reconnaissait les CRTs. Il avait déjà été montré que la calréticuline humaine pouvait interagir avec les GRC1q (**Païdassi et al., 2011**), alors qu'il a été décrit que la TcCRT interagissait plutôt avec les régions collagènes (**Ferreira et al., 2004a**). Pour décrypter ces interactions, les expériences ELISA ont été répétées avec les CLR et les GRC1q en incubant les différents fragments avec 10 nM de GRC1q et 10 nM ou 200 nM de CLR (concentration utilisée pour la TcCRT dans la littérature (**Ferreira et al., 2004a**)).

Nous avons pu constater qu'en utilisant seulement 10 nM de CLR (concentration utilisée pour C1q), nous n'observions aucune interaction. Nous nous sommes alors posé la question de savoir si cette différence pouvait être due à l'utilisation d'anticorps anti-C1q pour révéler les CLR. Un dosage effectué au laboratoire a mis en évidence que les anticorps anti-C1q reconnaissent C1q et les CLR de la même manière (résultats non montrés). *A contrario*, nous avons pu observer des interactions similaires à celles obtenues avec C1q en utilisant 10 nM de GRC1q. (**Figure 3.IV.3**).

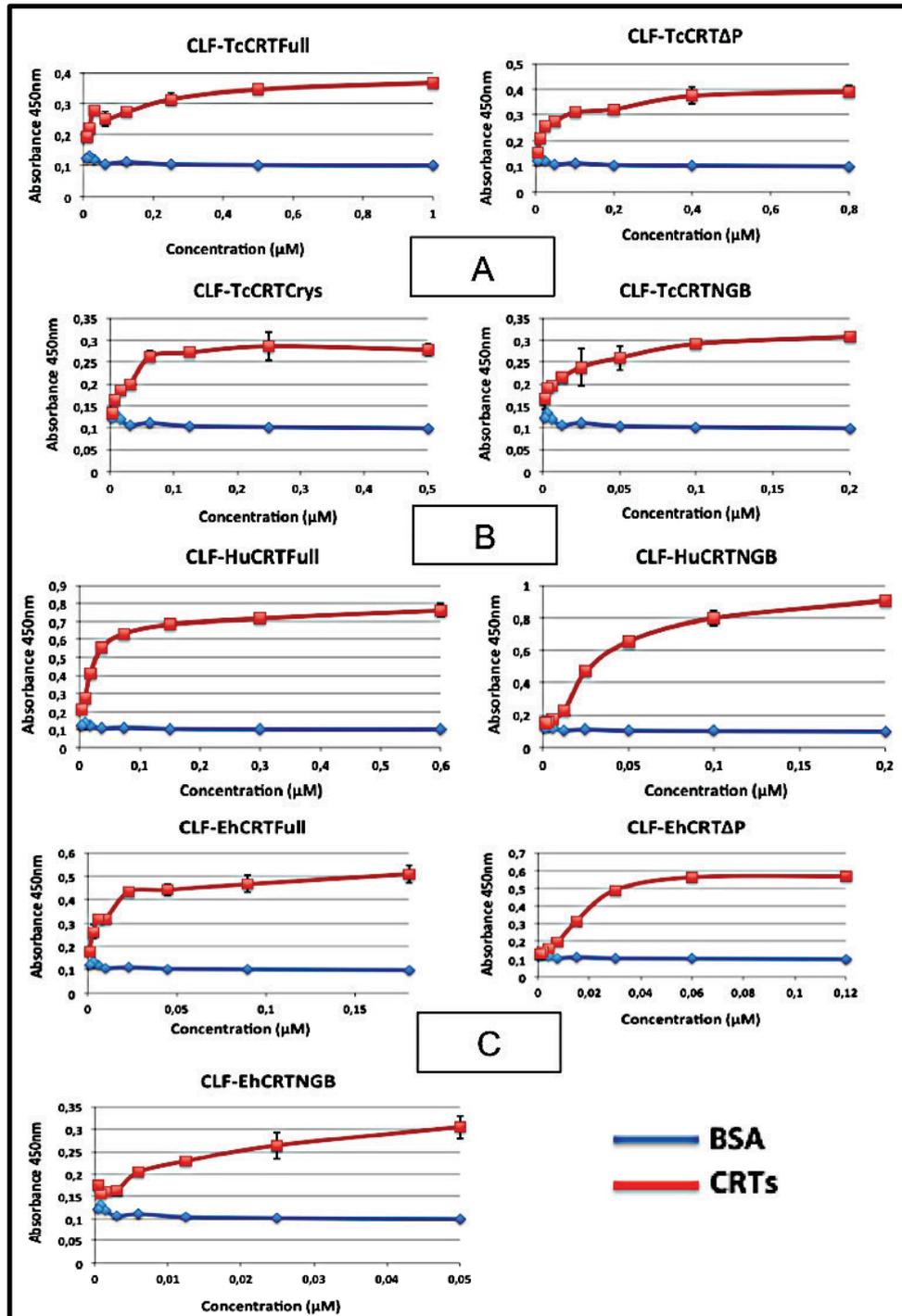


FIGURE 3.IV.2 : INTERACTIONS ENTRE LES CLR DE C1Q ET LES FRAGMENTS DE CALRETICULINES (ELISA)

Les différents fragments de calréticulines (Tc, Hu, Eh en A, B et C, respectivement) ou la BSA ont été adsorbés sur une surface solide à huit concentrations différentes et incubés avec les CLR de C1q à une concentration de 200 nM. L'interaction est révélée après lavage par l'ajout successif d'anticorps de lapin anti-C1q et d'anticorps dirigés contre les anticorps de lapin et couplés à la HRP. Pour que les résultats soient comparables, les différentes plaques ont été réalisées en parallèle et la réaction de coloration a été stoppée au même temps. Les courbes présentées sont le résultat de deux expériences indépendantes effectuées en triplicats (chaque point est la moyenne de 6 valeurs).

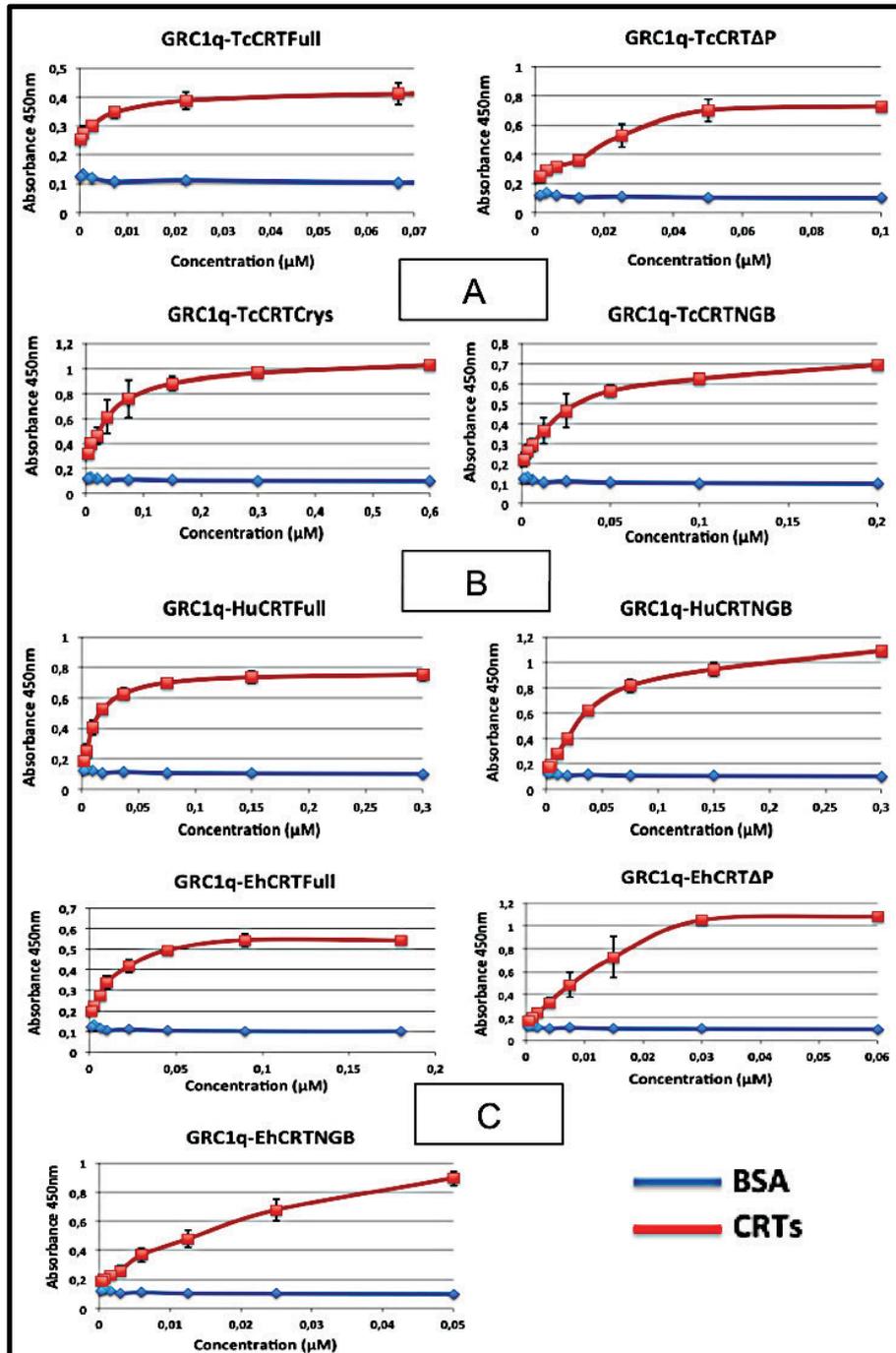


FIGURE 3.IV.3 : INTERACTIONS ENTRE LES GRC1Q ET LES FRAGMENTS DE CALRETICULINES (ELISA)

Les différents fragments de calreticulines (Tc, Hu, Eh en A, B et C, respectivement) ou la BSA ont été adsorbés sur une surface solide à huit concentrations différentes et incubés avec les GRC1q à une concentration de 10 nM. L'interaction est révélée après lavage par l'ajout successif d'anticorps de lapin anti-GRC1q et d'anticorps dirigés contre les anticorps de lapin et couplés à la HRP. Pour que les résultats soient comparables, différentes plaques ont été réalisées en parallèle et la réaction de coloration a été stoppée au même temps. Deux expériences indépendantes ont été effectuées en triplicat, donc chaque point est la moyenne de 6 valeurs. Nous avons vérifié que les anticorps secondaires ne reconnaissent pas les GRC1q.

A nouveau, les interactions avec les têtes globulaires sont plus fortes pour les fragments sans domaine P de la TcCRT et de la EhCRT (**Figure 3.IV.3 A et C**). Aucune différence n'est observée entre les fragments complets et sans domaine P pour la HuCRT (**Figure 3.IV.3 B**). Cette observation, associée à l'interaction forte de la HuCRT avec les CLR suggère une interaction du domaine P avec les CLR dans le cas de la protéine humaine ; cependant, en observant les interactions entre les CLR pour les différents fragments de la HuCRT, nous ne notons aucune différence en présence ou absence du domaine P (**Figure 3.IV.2 B**). Alors que l'absorbance à 450 nm mesurée pour la EhCRT et la TcCRT ne dépasse pas 0,5 dans ce cas, elle atteint facilement 1 pour la HuCRT. Il semblerait que la HuCRT interagisse plus fortement avec les régions collagènes que les protéines de parasite. Toutefois, même si l'interaction avec les domaines collagènes existe, nous avons du utiliser 200 nM contre 10 nM pour les GRC1q et C1q pour mettre en évidence une interaction (**Figure 3.IV.2**).

Ces résultats montrent que l'interaction entre les calréticulines et C1q s'effectue principalement *via* les domaines globulaires des deux protéines. De même, les trois protéines ne sont pas équivalentes, l'interaction est plus forte avec le ligand naturel HuCRT, puis avec la EhCRT et enfin la TcCRT est la protéine pour laquelle nous devons immobiliser la plus grande concentration pour atteindre le plateau.

Contrairement aux protéines de parasite, l'interaction entre C1q et la HuCRT impliquerait d'avantage les domaines collagènes de C1q, même s'ils ne représentent pas le facteur critique de la reconnaissance.

IV.2. Paramètres cinétiques de l'interaction entre C1q et les CRTs

Les expériences ELISA ont été essentiellement utilisées de manière qualitative. En effet, même si les expériences ont été réalisées en parallèle pour chaque fragment, de nombreux paramètres entrent en compte. La fixation de C1q et des GRC1q est par exemple révélée avec des anticorps primaires différents. L'absorbance lue sur la plaque n'est pas seulement liée à l'interaction entre les deux partenaires, mais aussi à l'affinité entre l'anticorps primaire et l'analyte, entre l'anticorps secondaire et l'anticorps primaire. Enfin nous avons observé que l'efficacité de la coloration était dépendante de paramètres indépendants des interactions (comme la température ou encore le lot de réactif TMB).

Pour confirmer les interactions observées et déterminer leurs paramètres cinétiques, ainsi qu'un potentiel modèle de fixation dans le meilleur des cas, nous avons réalisé des

expériences de SPR. Nous avons commencé par étudier l'interaction avec les GRC1q. Les différents fragments de CRTs ont été immobilisés en utilisant les conditions décrites pour immobiliser la CRT humaine (Païdassi et al., 2011). Pour effectuer une cinétique, cinq concentrations de GRC1q ont été successivement injectées (90 nM, 180 nM, 360 nM, 720 nM, 1,44 μ M) sans régénération (single cycle kinetics). Les sensorgrammes obtenus présentés sur la **figure 3.IV.4** confirment l'interaction de tous les fragments avec les GRC1q avec des affinités similaires à celles mesurées dans la littérature pour la HuCRT. Pour obtenir des informations cinétiques, les données ont été analysées en se basant sur un modèle d'interaction stoechiométrique 1:1. Ce modèle n'expliquait cependant pas la courbe biphasique observée, pour laquelle un modèle 1:1 impliquant un changement de conformation lors de l'interaction s'est avéré beaucoup plus adapté. Dans ce type de modèle, une première constante de vitesse (k_{a1}) est associée à la reconnaissance entre les partenaires pour former un complexe intermédiaire, puis une seconde (k_{a2}) correspond au changement de conformation permettant de stabiliser le complexe. Ensuite une constante de vitesse (k_{d2}) correspond au changement de conformation permettant de libérer le partenaire, alors que (k_{d1}) correspond à la dissociation des deux protéines.

Grâce aux paramètres cinétiques déterminés en utilisant ce modèle (**Tableau 3.IV.1**), nous observons une affinité K_D du même ordre (110 à 180 nM) pour tous les fragments des calréticulines des différentes espèces. Seul le fragment HuCRTNGB montre une affinité plus importante. Comme pour le fragment EhCRTNGB dont je discuterai plus tard, il est possible que la suppression du domaine P dévoile des régions hydrophobes supplémentaires, ce qui augmenterait la capacité d'interaction des fragments.

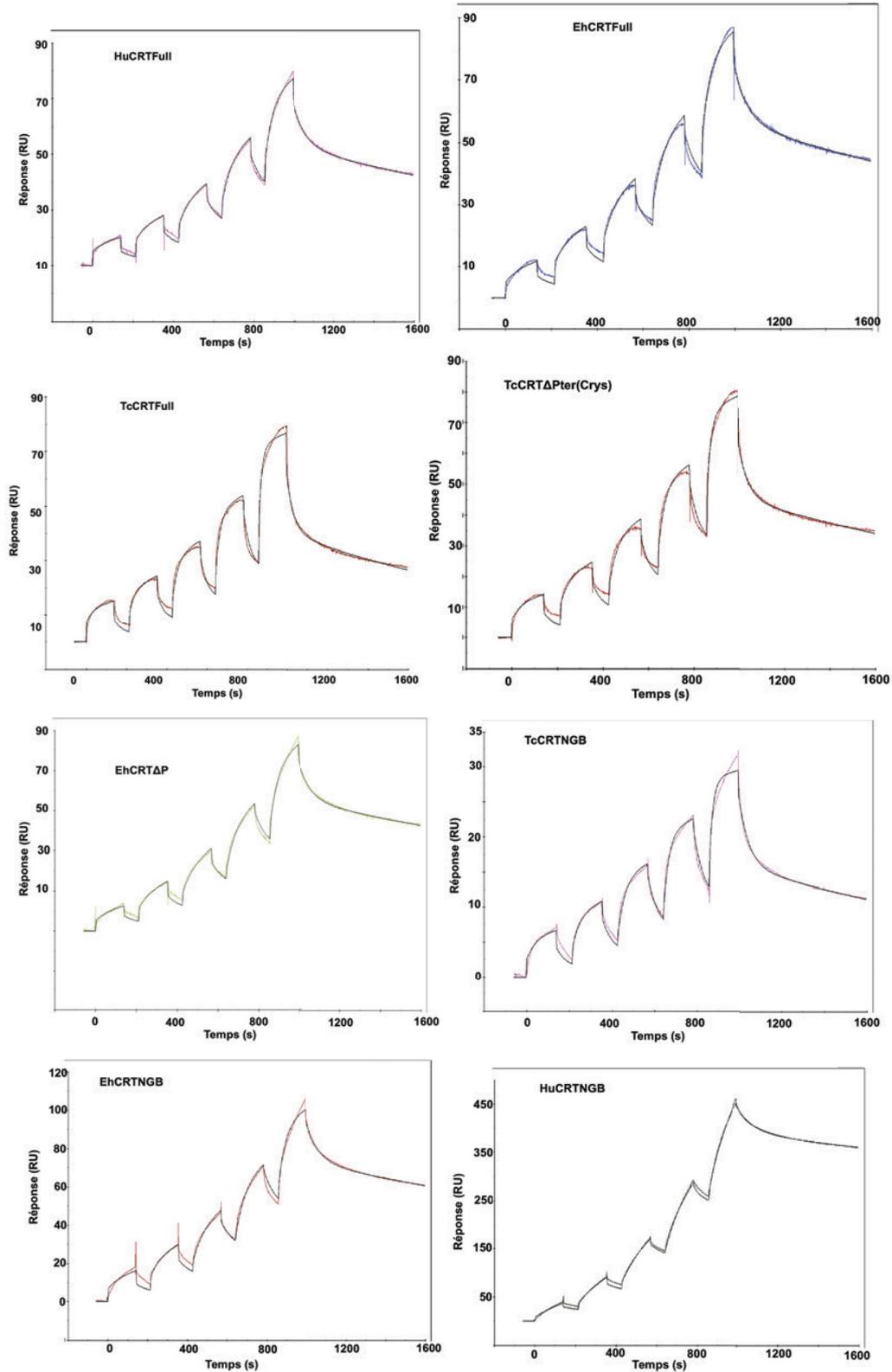


FIGURE 3.IV.4 : CINETIQUES D'INTERACTIONS ENTRE LES GRC1Q ET LES FRAGMENTS DE CRTS
 Les GRC1q ont été injectées à cinq concentrations (90 nM, 180 nM, 360 nM, 720 nM, 1,44 μ M) en tampon HBS-P sur les différents fragments de calréticulines immobilisés au niveau indiqué dans la section Matériels et Méthodes. Les courbes expérimentales sont présentées en couleur alors que les courbes ajustées sont en noir.

Les valeurs des constantes cinétiques analysées plus finement confirment une association et dissociation des protéines très similaires. En effet, les différentes constantes d'association et de dissociation sont quasi-identiques pour les fragments d'une même espèce et diffèrent au maximum d'un facteur 2 entre deux espèces, reflétant les constantes d'affinité. Les valeurs de Rmax observées montrent cependant que seules 5% des protéines immobilisées fixent les GRC1q (Rmax théoriques situées aux alentours de 3200 RU).

| Fragments de calréticulines (RU immobilisés) | GRC1q (modèle de fixation 1:1 avec un changement de conformation) | | | | | | |
|--|--|------------------|---|-----------------------------|-----------------------------|-----------------------------|------|
| | K_D (nM) | Chi ² | K_{a1} (M ⁻¹ s ⁻¹) | K_{a2} (s ⁻¹) | K_{d1} (s ⁻¹) | K_{d2} (s ⁻¹) | Rmax |
| HuCRTFull (3080) | 174 | 1,2 | 8,9.10 ³ | 4,5.10 ⁻³ | 1,2.10 ⁻² | 6,5.10 ⁻⁴ | 87 |
| HuCRTNGB (2357) | 53 | 15 | 4,8.10 ³ | 7,0.10 ⁻³ | 6,0.10 ⁻³ | 3,2.10 ⁻⁴ | 662 |
| TcCRTFull (2736) | 165 | 2,4 | 2,6.10 ⁴ | 3,8.10 ⁻³ | 2,3.10 ⁻² | 7,9.10 ⁻⁴ | 79 |
| TcCRTCrys (2037) | 113 | 2,0 | 2,2.10 ⁴ | 4,1.10 ⁻³ | 2,0.10 ⁻² | 5,8.10 ⁻⁴ | 80 |
| TcCRTNGB (2097) | 141 | 0,5 | 2,8.10 ⁴ | 3,8.10 ⁻³ | 1,8.10 ⁻² | 6,9.10 ⁻⁴ | 32 |
| EhCRTFull (2800) | 131 | 1,7 | 1,0.10 ⁴ | 4,9.10 ⁻³ | 1,3.10 ⁻² | 6,1.10 ⁻⁴ | 102 |
| EhCRTΔP (2210) | 124 | 2,3 | 9,1.10 ³ | 5,1.10 ⁻³ | 1,3.10 ⁻² | 5,2.10 ⁻⁴ | 121 |
| EhCRTNGB (2060) | 136 | 4,0 | 1,14.10 ⁴ | 4,0.10 ⁻³ | 1,1.10 ⁻² | 4,5.10 ⁻⁴ | 123 |

TABLEAU 3.IV.1 : CONSTANTES CINÉTIQUES DÉTERMINÉES PAR SPR POUR L'INTERACTION GRC1Q-CALRETICULINES EN SUIVANT UN MODÈLE D'INTERACTION 1:1 AVEC CHANGEMENT DE CONFORMATION

De la même manière, différentes concentrations de C1q ont été injectées sur les différents fragments (8 nM, 16 nM, 32 nM, 64 nM, 130 nM) en single cycle kinetics. Les différents sensorgrammes présentés **figure 3.IV.5** confirment à nouveau l'interaction avec tous les fragments. Toutefois, obtenir des paramètres cinétiques exploitables à l'aide d'un modèle de fixation prédéfini s'est révélé plus compliqué. Nous avons tout d'abord opté pour un modèle d'interaction 1:1. Ce modèle suppose que, une fois une molécule de calréticuline fixée à une

GRC1q, l'encombrement stérique empêche toute autre fixation de CRT, ou alors que C1q interagisse en se couchant. Comme pour les GRC1q, ce modèle étant moins adapté que le modèle 1 :1 incluant un changement de conformation. Il a été montré au laboratoire que l'interaction entre la HuCRT et C1q ne peut pas être observée dans la configuration inverse (C1q immobilisé et CRT soluble). Cette observation pourrait être expliquée par l'impossibilité de C1q à subir le changement de conformation nécessaire à l'interaction une fois immobilisé (**Païdassi et al., 2011**).

Les affinités mesurées à partir du modèle 1 :1 avec changement de conformation sont très proches pour les calréticulines de parasite et les données de la calréticuline humaine publiées (**Païdassi et al., 2011**). Par contre, les sensorgrammes obtenus pour les fragments humains et la EhCRT Δ P ne peuvent être analysés correctement, vraisemblablement du fait d'un niveau de fixation trop élevé (**Figure 3.IV.5**) pouvant favoriser des phénomènes d'interactions secondaires ou d'agrégation. L'expérience sera à reproduire en immobilisant moins de protéines.

Par ailleurs, les données obtenues pour le fragment EhCRTNGB sont assez étranges. En effet, le K_D mesuré de 2 nM correspond à une affinité extrêmement élevée en comparaison des autres fragments. Dans ce cas particulier, nous avons été obligés de diminuer la gamme de concentrations de C1q injectée pour obtenir des données exploitables. Il se pourrait que l'absence de domaine P favorise drastiquement l'interaction, soit par une diminution de l'encombrement stérique comme je l'ai déjà évoqué, mais cet effet ne se produisant pas ici pour la TcCRTNGB, il se pourrait également que l'absence du domaine P dévoile des régions hydrophobes favorisant les interactions non spécifiques *via* les régions collagènes (cet effet n'étant pas observé pour les GRC1q). Au contraire la fixation observée était assez faible sur la TcCRTNGB, potentiellement à cause des conditions de régénération mal supportées par ce fragment (**Figure 3.IV.5**).

Les valeurs de R_{max} montrent ici aussi que seuls 2 à 5% des sites de fixation étaient accessibles (le R_{max} théorique se situe aux alentours de 30000 RU).

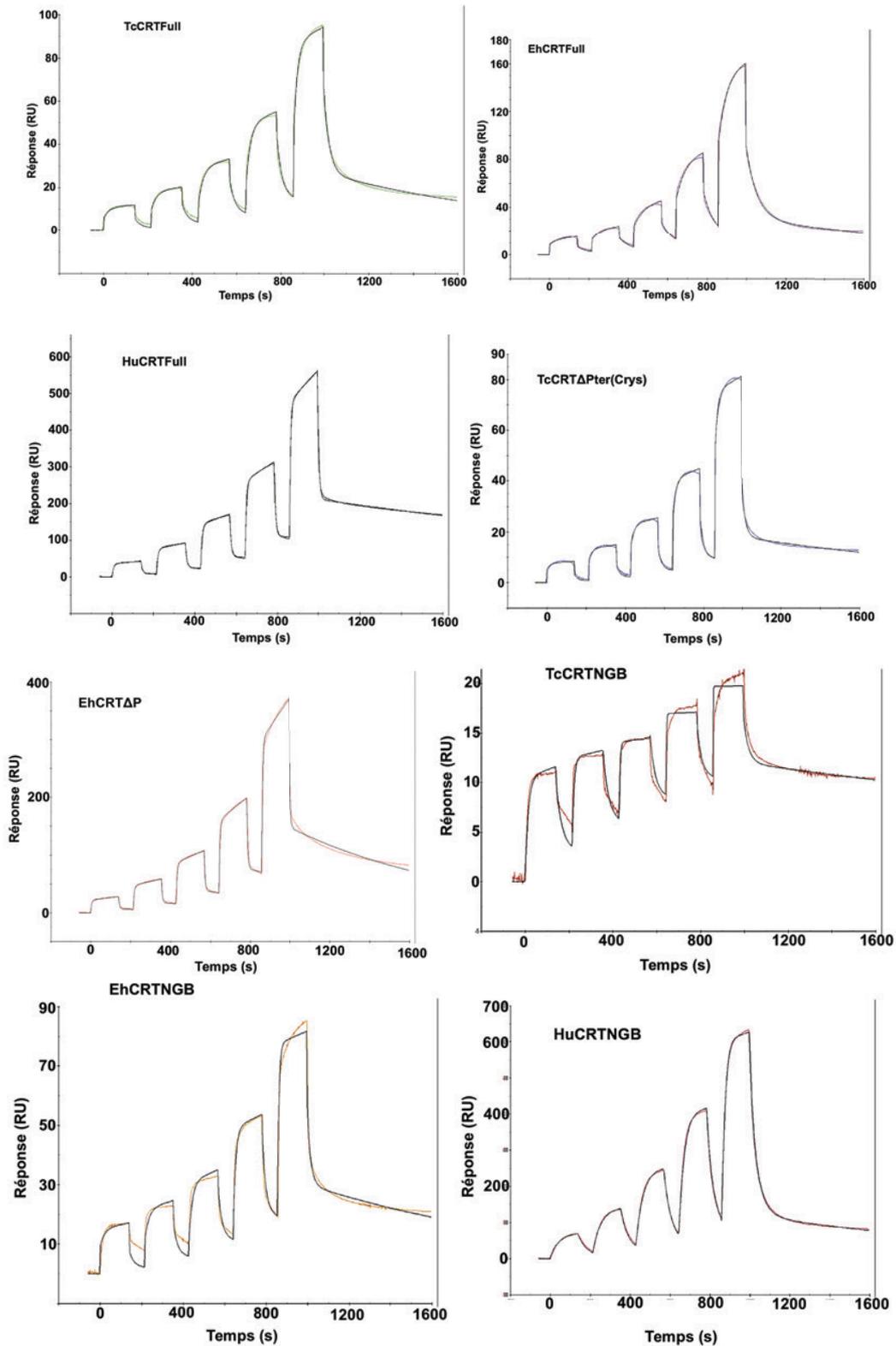


FIGURE 3.IV.5 : CINÉTIQUES D'INTERACTIONS ENTRE Clq ET LES FRAGMENTS DE CRTS

Clq a été injecté à cinq concentrations (8 nM, 16 nM, 32 nM, 64 nM, 130 nM) en tampon HBS-P sur les différents fragments de calréticulines immobilisés au niveau indiqué dans la section Matériels et Méthodes (ou à 2,5 nM, 5 nM, 10 nM, 20 nM, 40 nM sur les fragments EhCRTNGB et HuCRTNGB) (Single Cycle Kinetics). Les courbes expérimentales sont présentées en couleurs alors que les courbes ajustées sont en noires.

| Fragments de calréticulines (RU immobilisés) | C1q Modèle de fixation 1:1 avec changement de conformation | | | | | | |
|--|---|------------------|---|-----------------------------|-----------------------------|-----------------------------|------|
| | K_D (nm) | Chi ² | K_{a1} (M ⁻¹ s ⁻¹) | K_{a2} (s ⁻¹) | K_{d1} (s ⁻¹) | K_{d2} (s ⁻¹) | Rmax |
| TcCRTFull (2736) | 84 | 2,6 | 1,75.10 ⁵ | 2,1.10 ⁻³ | 4,2.10 ⁻² | 1,2.10 ⁻³ | 302 |
| TcCRTCrys (2037) | 84 | 1,04 | 1,6.10 ⁵ | 2,4.10 ⁻³ | 5,7.10 ⁻² | 7,07.10 ⁻³ | 124 |
| TcCRTNGB (2097) | 88 | 1,6 | 1,42.10 ⁴ | 2,5.10 ⁻³ | 5,3.10 ⁻² | 1,5.10 ⁻³ | 354 |
| EhCRTFull (2800) | 135 | 5,08 | 4,8.10 ⁴ | 1,6.10 ⁻³ | 1,7.10 ⁻² | 9,5.10 ⁻⁴ | 157 |
| EhCRTNGB (2060) | 2 | 4,0 | 1,14.10 ⁴ | 4,0.10 ⁻³ | 1,1.10 ⁻² | 4,5.10 ⁻⁴ | 90 |

TABLEAU 3.IV.2 : CONSTANTES CINÉTIQUES DÉTERMINÉES PAR SPR POUR L'INTERACTION C1Q-CALRÉTICULINES EN SUIVANT UN MODÈLE D'INTERACTION 1 :1 AVEC CHANGEMENT DE CONFORMATION

Avec ses six domaines globulaires, la protéine C1q peut théoriquement fixer jusqu'à six molécules de CRTs. Nous avons donc repris l'analyse des courbes en utilisant un modèle d'interaction bivalent pour nous rapprocher du modèle hexavalent théorique. Dans la littérature, les études sur les anticorps montrent bien que, en suivant un modèle bivalent, les anticorps complets possèdent une affinité et des constantes d'association supérieures aux fragments Fab d'un facteur 100 (Müller-Loennies et al., 2000). Cependant, dans notre cas, le choix de ce modèle ne modifie pas la valeur des paramètres mesurés et n'explique pas mieux la courbe cinétique.

Ces expériences ont confirmé que l'interaction entre les calréticulines et C1q s'effectuait majoritairement avec les régions globulaires de C1q et le domaine globulaire des CRTs, même s'il se pourrait que les régions collagènes puissent jouer un rôle pour la HuCRT (nous n'avons obtenu aucun résultat exploitable en SPR avec les CLR de C1q à cause d'interactions non spécifiques et une impossibilité à régénérer les surfaces). Néanmoins, il est très compliqué de définir un modèle d'interaction. Les valeurs très proches des constantes cinétiques et d'affinité obtenus pour C1q complet et les GRC1q suggèrent que chaque molécule de C1q n'aurait la possibilité d'interagir qu'avec une seule calréticuline. Ces observations semblent confirmées par le fait qu'un modèle bivalent ne modifie pas drastiquement les valeurs des constantes non plus.

IV.3. Importance du site lectine

Nous avons pu déterminer l'implication du domaine globulaire des calréticulines dans l'interaction avec C1q. Des modèles d'interaction ont été proposés dans la littérature, suggérant une surface d'interaction localisée sur la même face que le domaine lectine (Yadav et al., 2014). Nous nous sommes demandé si les acides aminés du domaine lectine pouvaient être impliqués dans l'interaction avec C1q. Pour cela, nous avons reproduit les expériences ELISA en immobilisant les fragments TcCRTFull et EhCRTFull et nous les avons incubés avec la protéine C1q seule ou en présence de 250 mM ou 500 mM de D-glucose. Les courbes d'interaction présentées sur la **figure 3.IV.6** montrent une inhibition dépendante de la concentration du glucose pour les deux protéines.

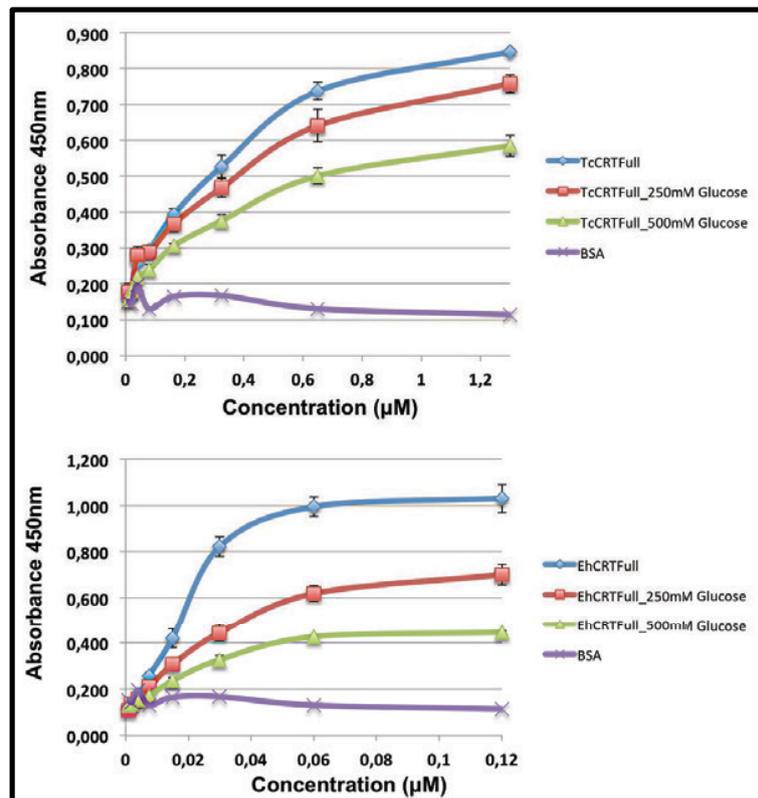


FIGURE 3.IV.6 : ESSAIS DE COMPÉTITIONS AVEC LE GLUCOSE

Les différents fragments de calréticulines ou la BSA ont été adsorbés sur une surface solide à huit concentrations différentes et incubés avec C1q seul à une concentration de 10 nM ou en présence de 250 mM ou 500 mM de D-glucose. L'interaction est révélée après lavage par l'ajout successif d'anticorps de lapin anti-C1q et d'anticorps dirigés contre les anticorps de lapin et couplés à la HRP. Pour que les résultats soient comparables, différentes plaques ont été réalisées en parallèle et la réaction de coloration a été stoppée au même temps. Les courbes présentées sont le résultat de deux expériences indépendantes effectuées en triplicat (chaque point est la moyenne de 6 valeurs).

Pour vérifier si réellement le glucose et C1q partageaient des sites d'interactions en commun ou si la présence de glucose avec C1q modifiait la viscosité de la solution et donc provoquait une inhibition en modifiant la thermodynamique de l'interaction et serait donc un artefact, nous avons reproduit l'expérience avec du mannose comme témoin positif. En effet, étant donné que dans le RE la suppression du glucose ne permet plus aux protéines de se fixer à la CRT malgré la présence de résidus mannose, ce dernier ne devrait pas se fixer à la CRT. Il a été vérifié au laboratoire que le mannose n'interagissait pas avec C1q. De manière inattendue, le mannose a favorisé l'interaction avec la TcCRT et la EhCRT. Ce résultat étonnant nécessite d'être confirmé, mais laisse néanmoins supposer que l'inhibition observée était bien due au glucose (**Figure 3.IV.7**)

J'ai introduit au début du manuscrit deux mécanismes potentiels d'interaction avec les protéines glucosylées ou non glucosylées (section II.2) (**Wijeyesakere et al., 2013**). Dans le second cas, il a été proposé que le domaine P interviendrait dans l'interaction en se refermant, alors que la présence du glucose le maintiendrait en position ouverte. Nous nous sommes donc demandé si l'inhibition était due à un site d'interaction partagé entre C1q et le glucose, ou si le glucose maintenait le domaine P en position ouverte et défavorisait alors l'interaction avec C1q. Pour cela nous avons répété l'expérience avec le glucose sur les fragments NGB. En effet, si l'inhibition dépend du domaine P, le glucose ne devrait avoir aucun effet sur ces fragments. Comme montré sur la figure **3.IV.7**, le glucose inhibe également l'interaction des fragments NGB, ce qui implique que l'inhibition pourrait être associée à une compétition.

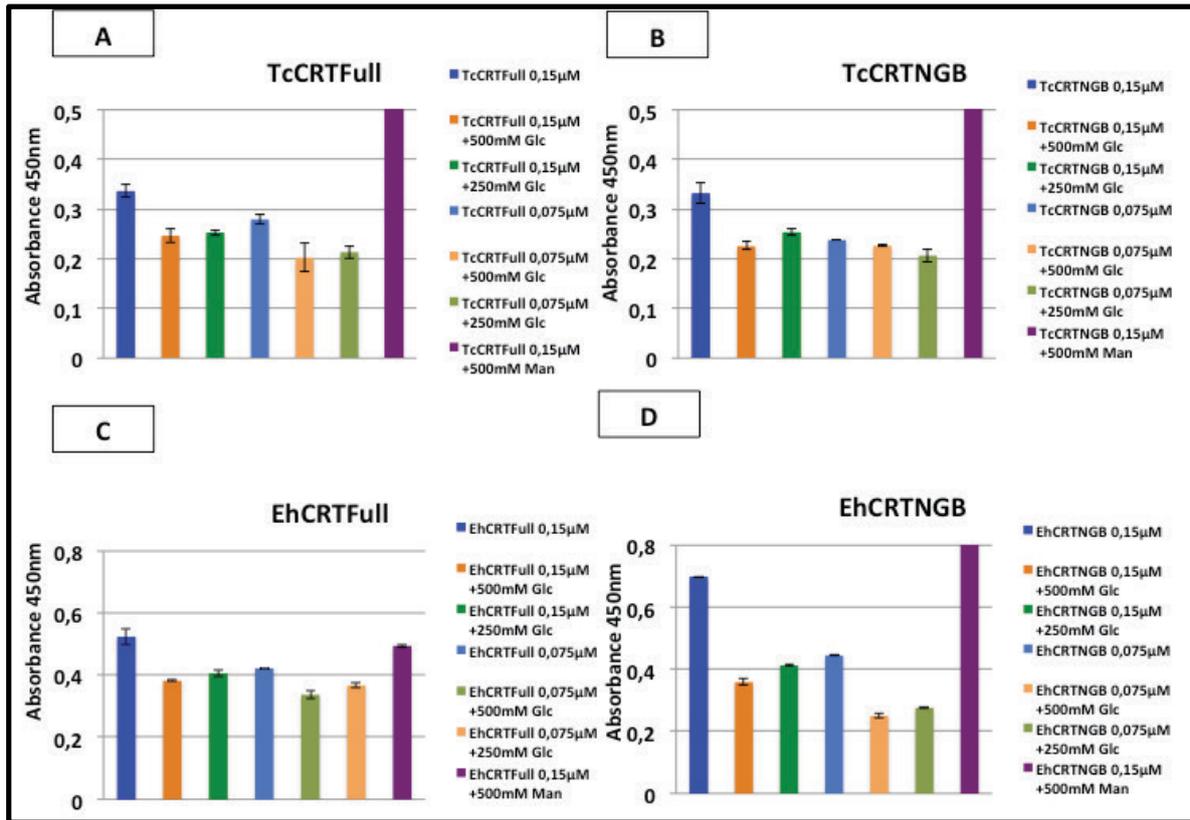


FIGURE 3.IV.7 : EFFETS DU GLUCOSE ET DU MANNOSE SUR L'INTERACTION C1q-CRT

La TcCRT Full (A), la TcCRTNGB (B), La EhCRTFull (C) et la EhCRTNGB (D) ont été adsorbées sur une surface solide à 0,15 et 0,75 μM et incubées avec C1q seul à 10 nM ou en présence pour chaque concentration de 250 mM ou 500 mM de glucose ou de 500 mM de mannose comme contrôle positif. L'interaction est révélée après lavage par l'ajout successif d'anticorps de lapin anti-C1q et d'anticorps dirigés contre les anticorps de lapin et couplés à la HRP. Les valeurs ne reflètent qu'une expérience réalisée en triplicata. Le témoin mannose est représenté en violet dans chacun des histogrammes.

IV.4. Vers une identification des acides aminés clés impliqués dans l'interaction

Au cours des dernières années, des modèles d'interaction entre les CRT et C1q ont été proposés (Yadav et al., 2014) et des sites potentiels d'interaction ont été proposés. Au sein de notre laboratoire, plusieurs mutations que je discuterais à la fin de ce manuscrit, ont été effectuées sur la HuCRT, qui se sont avérées sans effet sur l'interaction avec C1q. Récemment la production de la première forme recombinante de C1q (Bally et al., 2013) nous a offert la possibilité d'étudier par mutagenèse les acides aminés de C1q impliqués dans ces interactions. Nous avons reproduit les expériences SPR avec la HuCRT entière et les fragments NGB des deux calréticulines de parasite, en injectant deux mutants de la protéine C1q recombinante entière, produits et purifiés au laboratoire par Isabelle Bally, et comparé

leurs effets par rapport à la protéine C1q recombinante sauvage. Ces mutants ont été sélectionnés pour inhiber l'interaction de C1q avec les immunoglobulines (mutation des R108 et R109 en D sur la chaîne B) ou l'héparine (mutation des K200 et K201 en D et E sur la chaîne A). Les premiers tests d'interaction de ces mutants (réalisés au laboratoire par Isabelle Bally) montre qu'ils sont toujours capables d'interagir avec les protéases MASPs et le tétramère C1r2/C1s2. De plus le mutant K200 et K201 peut se fixer aux IgG et IgM et restaurer l'activation de la voie classique du complément (données non présentées)

Le remplacement de deux charges positives par des charges négatives inhibe totalement l'interaction avec les trois calréticulines, comme illustré pour la HuCRTFull (**Figure 3.IV.8**). Le fait que la modification de seulement deux charges abolisse totalement l'interaction suggère une interaction de type électrostatique avec des résidus acides de la CRT.

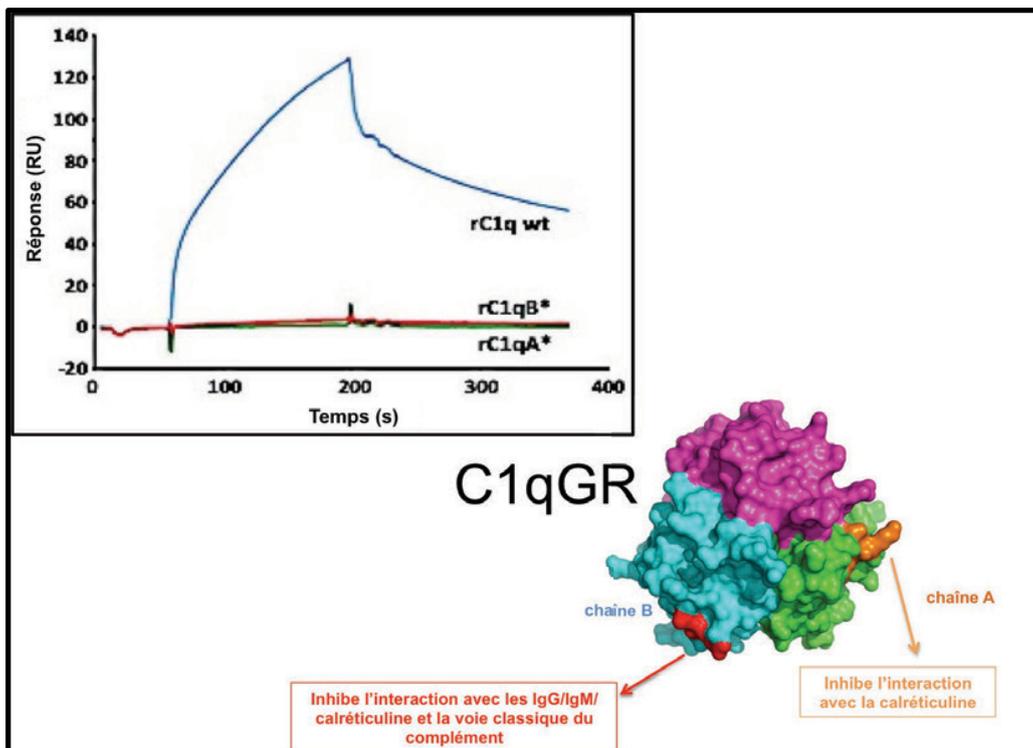


FIGURE 3.IV.8 : RESIDUS DE C1Q MODULANT L'INTERACTION AVEC LES CALRETICULINES

La forme recombinante de C1q sauvage et les mutants R108-109^BDD et K200-201^ADE ont été injectés à 10 nM en tampon HBS-P sur la HuCRTFull, TcCRTNGB et EhCRTNGB (immobilisées au niveau indiqué dans la section Matériels et Méthodes). Le sensorgramme présenté correspond à la HuCRTFull.

IV.5. Vers l'analyse structurale de l'interaction entre la TcCRT et les GRC1q par RMN

Dans le but de définir précisément les acides aminés essentiels à l'interaction entre les calréticulines et les GRC1q, nous avons pour objectif initial d'analyser le complexe à un niveau moléculaire. Dans la première partie de ma thèse, j'ai tout d'abord essayé de co-cristalliser la calréticuline de *T. cruzi* et les domaines globulaires de C1q, purifiés à partir du plasma, dans le but de résoudre leur structure en complexe. Malheureusement, je n'ai pas réussi à obtenir de cristaux du complexe. Sur la fin de la thèse, j'ai donc décidé de me tourner vers la RMN, et de réaliser des tests préliminaires pour voir si cette méthode expérimentale pourrait nous apporter des informations précises quant aux acides aminés de la calréticuline impliqués dans l'interaction.

IV.5.a Caractérisation de la TcCRT-N¹⁵

Afin de caractériser la TcCRT et observer les interactions en solution, cette dernière a tout d'abord été marquée avec de l'azote 15, comme décrit dans la section Matériels et Méthodes. Après avoir vérifié que la protéine restait soluble suite à une production en milieu minimum, nous l'avons produite et purifiée en utilisant le protocole établi. La croissance bactérienne étant plus lente en milieu minimum, nous avons augmenté la température de la phase de croissance avant induction de 30 à 37°C. Les spectres RMN ont été ensuite enregistrés et traités par Adrien Favier du groupe RMN.

Lors de la première étape, une expérience ¹H-¹⁵N-TROSY a été enregistrée sur la TcCRTΔPterΔ288 seule, à une concentration de 380 μM à 25°C. Cette expérience permet de corréler le proton et l'azote des liaisons amides. Le spectre enregistré et présenté **figure 3.IV.8**, nous a permis d'effectuer plusieurs contrôles. Nous avons confirmé que la protéine était repliée du fait de la dispersion des déplacements chimiques des NH, associée à un environnement varié des chaînes latérales. Comme le montre la **figure 3.IV.8**, certains de ces NH présentent des déplacements chimiques correspondant à des feuillets β (δH entre 9 et 10 ppm) et d'autres à des régions random-coil (δH entre 8 et 8,5 ppm), ce qui est en accord avec le cœur en feuillet β et les nombreuses boucles de la protéine observées dans la structure. Entre 10 et 10,5 ppm nous pouvons observer le déplacement chimique caractéristique des NH

des tryptophanes, dont le nombre correspond bien au nombre de tryptophanes présents dans le fragment (5).

Cette expérience nous a confirmé que la protéine était compatible avec une expérience de RMN, pour l'attribution de chacun des pics (taille, flexibilité, choix du tampon) au vu de la dispersion et de l'intensité des déplacements chimiques. Suite à ces observations et à la qualité du signal, il a été jugé possible de travailler à une concentration de 40 μM de protéine, en diminuant la concentration en sels de 150 à 100 mM dans le tampon. L'échange de protons entre la protéine et le solvant fait disparaître leur signal RMN et cet échange s'effectue d'autant plus rapidement que le pH est élevé. De ce fait, nous avons décidé pour la suite des expériences de diminuer le pH de 8 à 7 pour limiter ce phénomène. Nous avons testé également la possibilité d'enregistrer un spectre à 37°C, ce qui améliore souvent le rapport signal/bruit. Comme la TcCRT précipite partiellement à cette température, nous avons maintenu les essais suivants à 25°C.

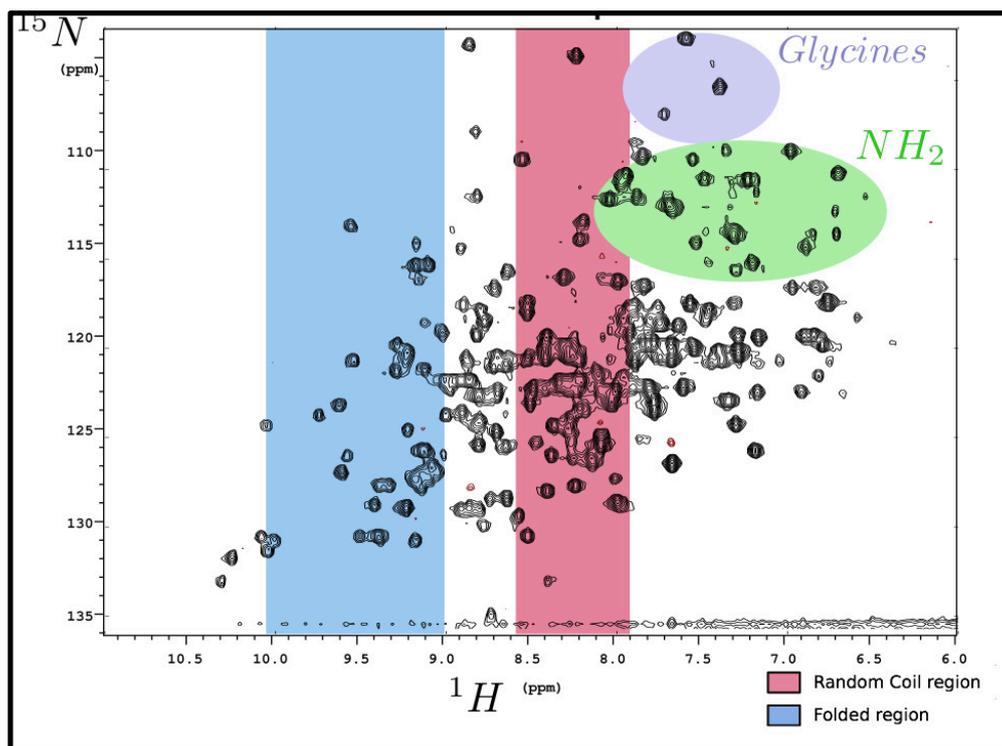


FIGURE 3.IV.8: CARACTERISATION DE LA TcCRT Δ PTER Δ 288 PAR UNE EXPERIENCE RMN HSQC ^{15}N - ^1H

IV.5.b Etudes préliminaires du complexe TcCRT-GRC1q

Pour que l'étude d'un complexe dont le partenaire non marqué possède une masse supérieure à environ 50 kDa soit possible en RMN, il est préférable que l'affinité soit faible (mM, μ M), afin que l'association et la dissociation du complexe soient rapides en solution. Dans le cas d'une interaction forte, les deux protéines restent associées et se comportent alors comme une particule de taille très importante, ce qui va conduire à l'élargissement des pics et à la perte du signal observé par RMN. Dans un équilibre d'association/dissociation plus rapide, il sera possible d'observer le déplacement chimique entre la forme libre et la forme en complexe des acides aminés participant à l'interaction. Dans ce cas, les déplacements chimiques des NH dont l'environnement électronique va être modifié par la présence du ligand non marqué vont progressivement changer des valeurs correspondant à la forme libre avant l'ajout des valeurs correspondant à la forme liée après avoir saturé la protéine marquée avec le ligand.

Nous avons donc réalisé des premiers essais en effectuant une expérience HSQC ^{15}N - ^1H sur la TcCRT seule à une concentration de 300 μ M et la forme en complexe avec 40 μ M de GRC1q (stoechiométrie 7:1). Nous avons pu observer l'apparition de certains déplacements chimiques pour la forme en complexe.

Devant ces résultats, nous avons décidé de réaliser une expérience de titration en enregistrant des spectres sur différents mélanges de protéines. Face à notre incapacité à concentrer de manière stable sur le long terme les GRC1q à une concentration supérieure à 40-50 μ M, nous avons travaillé avec une concentration maximale de TcCRT- N^{15} à 40 μ M pour être en mesure de réaliser un mélange stoechiométrique. Pour pallier la diminution du signal liée à la plus faible concentration et pour améliorer le rapport signal/bruit, les spectres HSQC ont été enregistrés sur des durées de 3h au lieu de 20 minutes. Les quatre solutions de protéines (5 μ M GRC1q-40 μ M TcCRT, 10 μ M GRC1q-40 μ M TcCRT, 20 μ M GRC1q-40 μ M TcCRT et 40 μ M GRC1q-40 μ M TcCRT) ont été dialysées dans un tampon unique 20 mM Tris, 100 mM NaCl, 5 mM CaCl_2 , pH 7.

Les spectres obtenus ont été superposés, et sont présentés sur la **figure 3.IV.9**. Nous constatons que, même dans le cas du mélange équimolaire, nous n'éliminons pas tous les pics de la forme libre. Cela signifie que l'interaction n'est pas aussi forte que nous le pensions et la totalité de la protéine marquée n'est pas sous forme de complexe. Cependant, alors que nous nous attendions à observer des déplacements de pics et leurs disparitions, nous avons observé

l'apparition de certains pics exclusifs à la forme complexée, ce qui signifie que nous sommes dans le cas de figure d'un échange lent à l'échelle de la RMN.

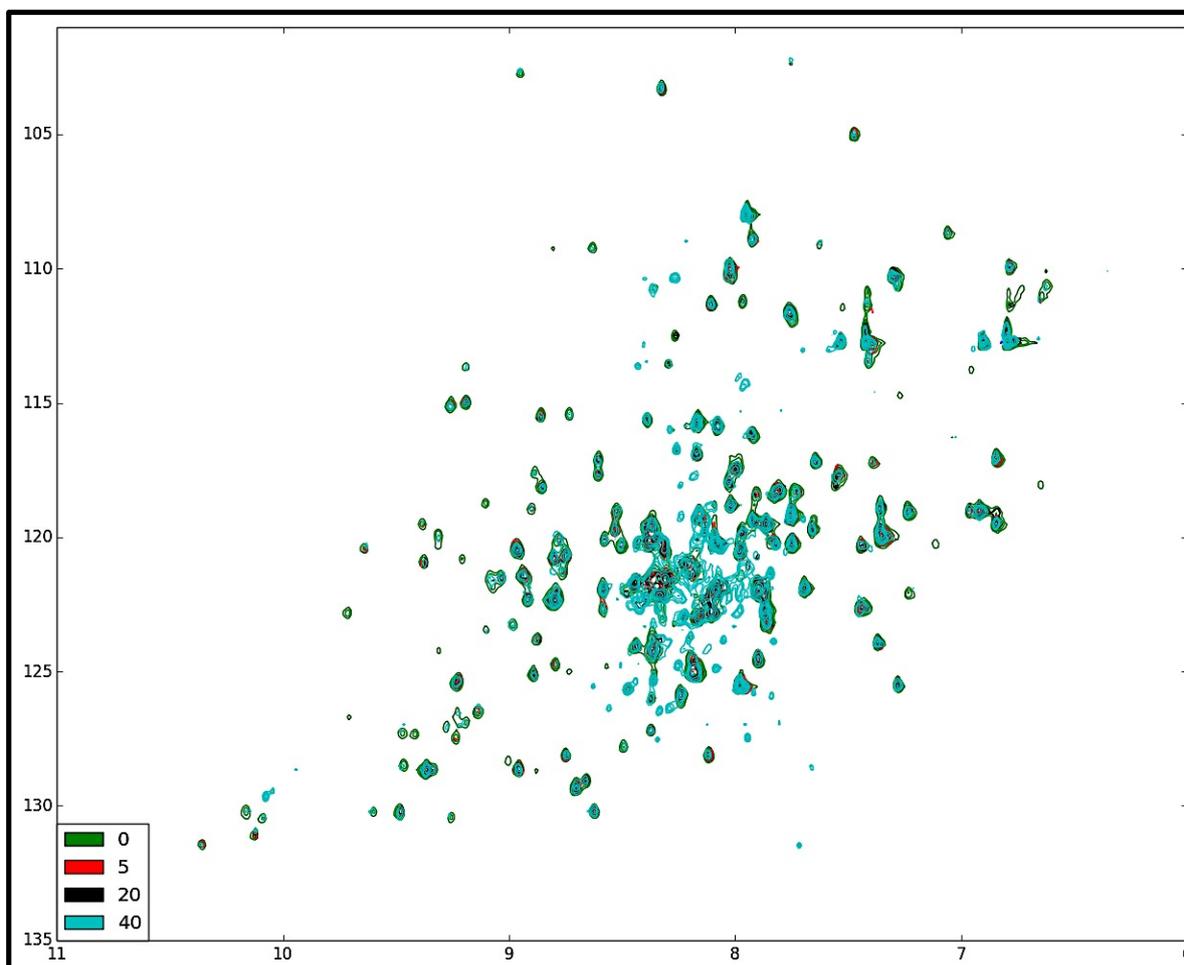


FIGURE 3.IV.9: SPECTRE RMN HSQC N¹⁵-H DE LA TcCRTΔPTerΔ288 EN COMPLEXE AVEC DIFFERENTES CONCENTRATIONS DES GRC1Q

Les nombreux pics apparus dans la forme complexée à équimolarité se situent entre 8 et 8,5 ppm, c'est à dire dans la zone du spectre correspondant aux régions random-coil (**Figure 3.IV.10**). Par ailleurs, ces pics étant très fins et résolus, ils pourraient être associés à de petites protéines ou peptides dynamiques en solution et pourraient donc correspondre à une dégradation de la protéine marquée. Afin de tester cette hypothèse, nous avons prévu d'enregistrer à nouveau un spectre sur la forme libre de la protéine et vérifier si elle s'est dégradée au cours du temps.

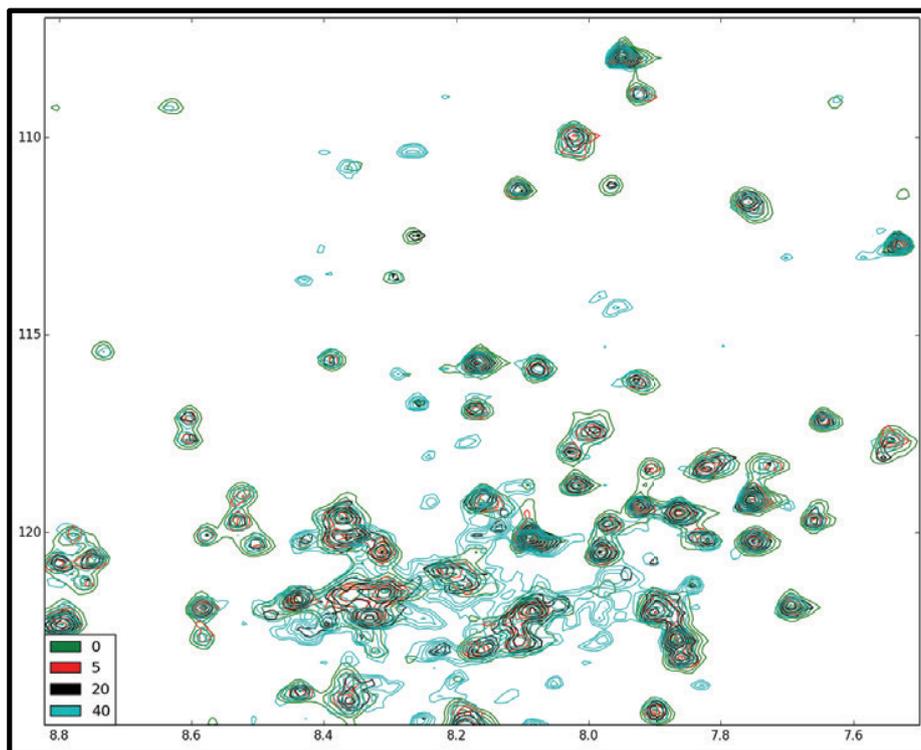


FIGURE 3.IV.10: AGRANDISSEMENT DE LA REGION RANDOM-COIL DU SPECTRE RMN HSQC TROSY ^{15}N - ^1H DE LA TcCRT Δ PTER Δ 288 EN COMPLEXE AVEC DIFFERENTES CONCENTRATIONS DES GRC1Q PLASMATIQUES

S'il s'avère que la protéine est intacte, il se pourrait alors que la formation du complexe augmente la flexibilité de certaines régions de la TcCRT et déstructure partiellement la protéine.

Parmi les différents pics apparus, l'un d'eux est très intéressant puisqu'il se situe dans la région du spectre spécifique des tryptophanes (**Figure 3.IV.11**). Ce déplacement chimique spécifique du NH de la chaîne latérale des tryptophanes dynamiques indique que l'un de ces résidus lors de la formation du complexe s'exposerait au solvant et deviendrait très dynamique.

Le fait que de nombreux pics apparaissent sans disparition associée de pics dans la forme libre ne nous permet pas d'identifier les résidus correspondant en réalisant une attribution sur la forme libre. Dans ce contexte, nous sommes obligés d'effectuer une attribution de la forme complexée de la protéine et l'instabilité des GRC1q à haute concentration nous empêche de réaliser ce type d'expériences. Néanmoins, comme je l'ai déjà abordé, nous pouvons observer de nombreux pics présents uniquement dans la forme libre (**Figure 3.IV.12**). En réalisant une titration inverse, c'est à dire en ajoutant un excès de protéine non marquée, nous pourrions saturer la protéine marquée et faire disparaître les pics impliqués dans l'interaction. Ensuite,

en attribuant les résonances restantes, nous pourrions déterminer par déduction les acides aminés les plus fortement impliqués dans la formation du complexe.

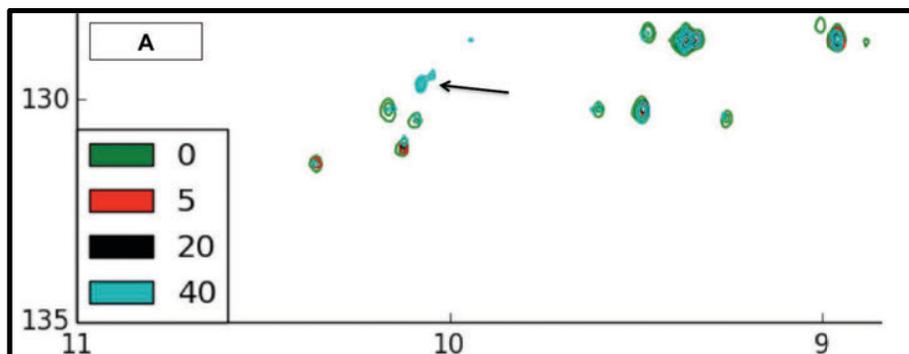


FIGURE 3.IV.11: PARTICIPATION DES TRYPTOPHANES AU COMPLEXE ENTRE LA TcCRT Δ Pter Δ 288 ET LES GRC1Q

(A) L'agrandissement de la région du spectre HSQC N15-H de la TcCRT Δ Pter Δ 288 correspondant aux NH des tryptophanes montre l'apparition d'un pic aux alentours de 10 ppm (flèche noire) pour la forme complexée, correspondant au NH de la chaîne latérale d'un tryptophane random-coil.

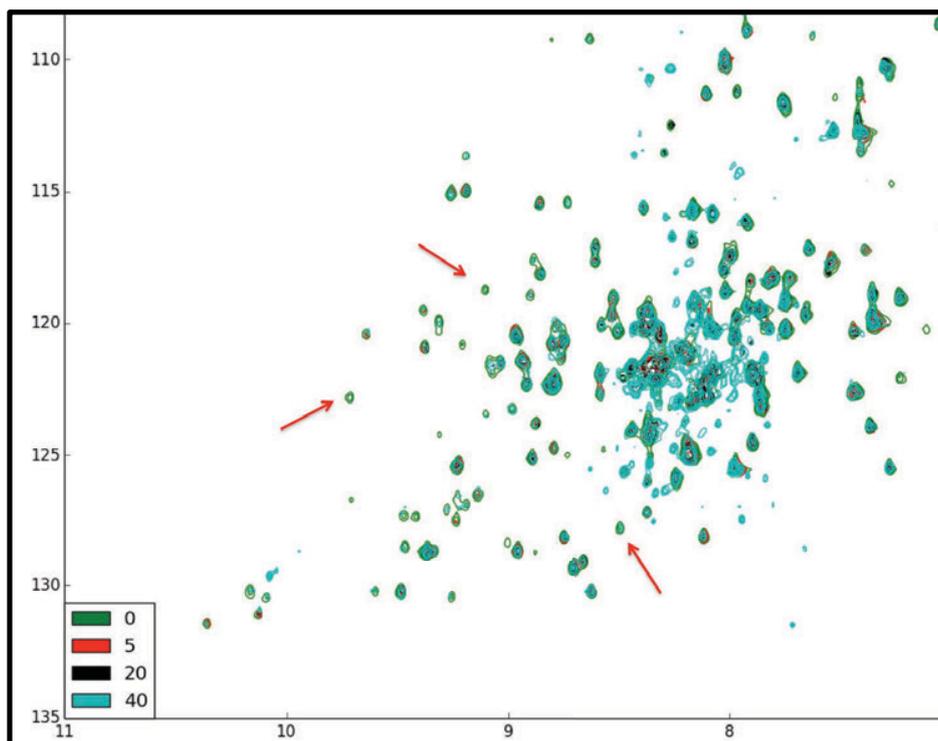


FIGURE 3.IV.12: RESIDUS PRINCIPAUX POTENTIELLEMENT IMPLIQUES DANS LE COMPLEXE ENTRE LA TcCRT Δ Pter Δ 288 ET LES GRC1Q

Certains déplacements chimiques sont présents uniquement dans la forme libre de la TcCRT (flèches rouges). Cela signifie que les déplacements chimiques correspondant à ces résidus se sont soit fortement déplacés ou ont totalement disparu dans la forme en complexe. Les résidus associés doivent être fortement impliqués dans le complexe.

Chapitre 3 : Développement de la première forme recombinante des régions globulaires de C1q

Pour réaliser les différentes études d'interaction, ou les analyses structurales, nous avons produit les GRC1q grâce à une digestion par la collagénase de la protéine C1q purifiée à partir du plasma humain. Les GRC1q sont ensuite purifiées par chromatographie d'exclusion de taille, qui nous permet d'obtenir 5 à 10 mg de protéine à partir de 2 L de plasma après 2 semaines de purification. De plus, travailler avec une protéine plasmatique ne permet pas d'étudier les résidus clés impliqués dans des interactions par mutagenèse dirigée. Récemment, la production au laboratoire par Isabelle Bally de la première forme recombinante de C1q produite en cellules de mammifères (**Bally et al., 2013**) résout ce problème en nous permettant de substituer les acides aminés d'intérêt. Néanmoins, l'instabilité de cette protéine à haute concentration et sa flexibilité pose des problèmes pour les études structurales en utilisant des techniques telles que la RMN, le SAXS et la cristallographie aux rayons X. Nous avons donc décidé de mettre au point et de produire en cellules de mammifère une forme recombinante des GRC1q dans laquelle les trois chaînes sont reliées par des linkers.

V. Construction du plasmide d'expression pcDNA3.1-GRC1q

Comme indiqué dans la section Matériels et Méthodes, un fragment codant pour les chaînes B et C et les deux linkers (linker A-C : GSGKQ et linker C-B : GSGKA) a été amplifié par mutagenèse dirigée à partir du gène synthétique présent dans le plasmide pNTBac et inséré dans le plasmide pcDNA3.1, contenant la séquence de la chaîne A des GRC1q précédé du peptide signal de C1qA. La séquence de la construction et un alignement avec les GRC1q plasmatiques sont présentés dans l'annexe 3, page C.

L'analyse du produit de PCR par électrophorèse sur gel d'agarose 1% a permis de montrer l'amplification d'un unique fragment d'environ 1000 paires de bases, ce qui correspond à la taille attendue de la séquence d'intérêt (835 pb) (**Figure 3.V.1 A**). Ce fragment a ensuite été purifié pour être utilisé dans l'étape de mutagenèse dirigée. Une fois sélectionnés et amplifiés, les plasmides néo-synthétisés issus de la réaction de mutagenèse ont été digérés par des enzymes de restriction possédant des sites de coupure situés de part et d'autre de l'insert. Le produit de digestion a été analysé sur gel d'agarose 1% en utilisant le plasmide sauvage digéré

par les mêmes enzymes comme témoin (**Figure 3.V.1 B**). Le séquençage (non présenté) du plasmide numéro 10 a confirmé l'insertion correcte de notre séquence d'intérêt au sein du plasmide pcDNA3.1

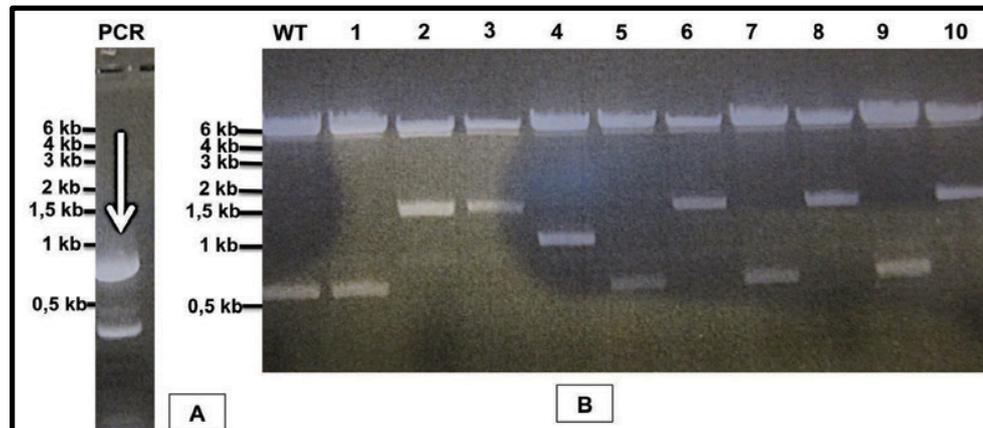


FIGURE 3.V.1: CLONAGE DE LA SEQUENCE CODANT LES GRC1Q RECOMBINANTES DANS LE PLASMIDE PCDNA3.1

(A) Le produit d'amplification par PCR migre sur gel d'agarose 1% à une distance correspondant à une taille apparente de 1000 pb correspondant à la séquence d'intérêt (835 pb) (flèche blanche). (B) La digestion du plasmide sauvage (WT) par les enzymes de restriction EcoRI et SacI génère un fragment de 557 paires de bases, observé également pour les plasmides 1, 5, 7 et 9. La présence de l'insert génère un fragment de 1362 paires de bases comme observé dans les plasmides 2, 3, 6, 8 et 10.

VI. Expression en cellules de mammifères

Le plasmide pcDNA3.1 contenant le gène codant pour les GRC1q recombinantes a été amplifié, purifié et utilisé pour transfecter des cellules humaines HEK 293-F. Après 72 h, la production des GRC1q a été contrôlée par Western Blot grâce à des anticorps anti-GRC1q révélant une bande unique migrant à une distance correspondant à la masse attendue pour les GRC1q monocaténaires (48 kDa), au contraire des GRC1q plasmatiques utilisées comme témoin, pour lesquelles les trois chaînes sont séparées (**Figure 3.VI.2**). Une fois le gène des GRC1q intégré au génome des cellules HEK 293-F grâce à la pression de sélection exercée par la néomycine, le surnageant de milieu de culture a été prélevé toutes les 72 h, jusqu'à obtenir un volume de 1,1 L. La production des GRC1q recombinantes a été à nouveau confirmée dans le surnageant final avant purification (**Figure 3.VI.3**). la quantité de protéine a été suffisante pour observer la protéine par coloration directe du gel.

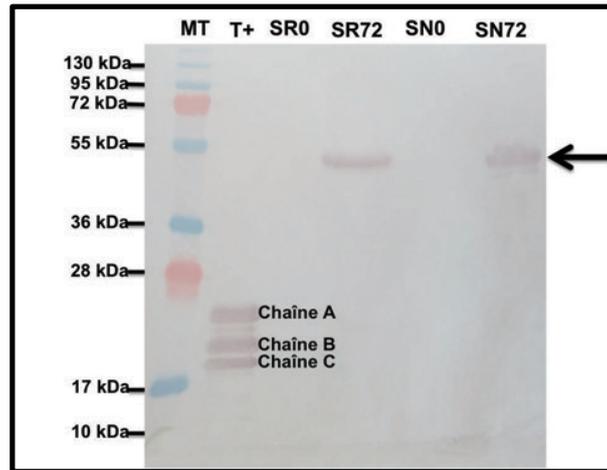


FIGURE 3.VI.2: TEST D'EXPRESSION DES GRC1Q RECOMBINANTES EN CELLULES HEK 293-F

Après 72 h, les protéines contenues dans 500 μ L de surnageant de milieu de culture sont précipitées au TCA 50%, séparées sur gel SDS-PAGE 12,5% et transférées sur une membrane de nitrocellulose. La présence des GRC1q recombinantes est révélée sur la membrane de nitrocellulose grâce à un anticorps anti-GRC1q. Les GRC1q plasmatiques sont utilisées comme témoins (MT : Marqueur de Taille, T+ : GRC1q plasmatique, SR0 et 72 : Surnageant Réduit prélevé après 0 et 72 h, SN0 et 72 : Surnageant Non Réduit prélevé après 0 et 72 h).

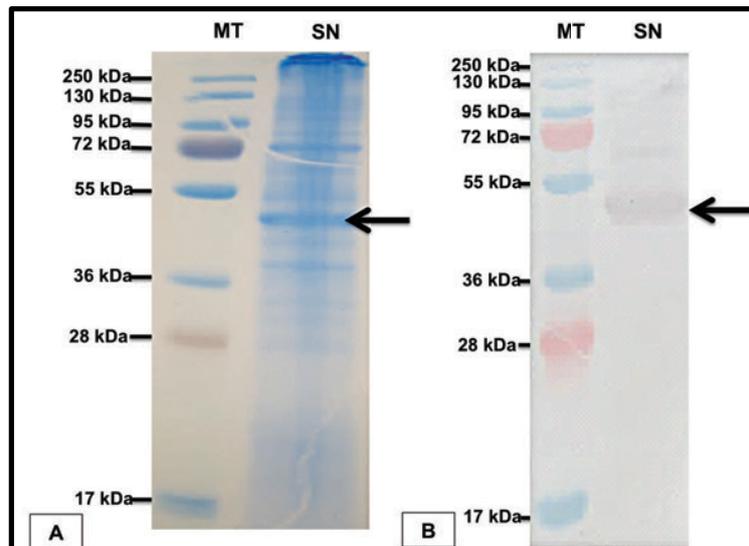


FIGURE 3.VI.3: CONTROLE DE LA PRODUCTION DES GRC1Q RECOMBINANTES PAR LES CELLULES HEK 293-F TRANSFECTEES DE MANIERE STABLE

Après 2 semaines de culture, les protéines contenues dans 500 μ L de surnageant de milieu de culture sont précipitées au TCA50% et déposées sur gel SDS-PAGE 12,5% (A) ou transférées sur une membrane de nitrocellulose (B). Une bande migrant à une distance correspondant à la masse apparente de la protéine d'intérêt (48 kDa) est visible sur le gel et révélée sur la membrane de nitrocellulose grâce à un anticorps anti GRC1q (MT : Marqueur de Taille, SN : Surnageant Non réduit).

VII. Purification des GRC1q recombinantes

Les GRC1q ont été purifiées en une étape unique de chromatographie échangeuse de cations. Un volume de 500 mL de surnageant de milieu de culture a été déposé sur une colonne SP-Sépharose et l'élution est effectuée par un gradient de NaCl allant de 25 mM à 500 mM (2x500 mL). La mesure de l'absorbance à 280 nm a permis de tracer le profil d'élution (**Figure 3.VII.1**) mettant en évidence un pic majeur élué à un volume correspondant à 160 mM de NaCl. La présence des GRC1q recombinantes en grande quantité a été confirmée par analyse sur gel SDS-PAGE et Western Blot (non présenté).

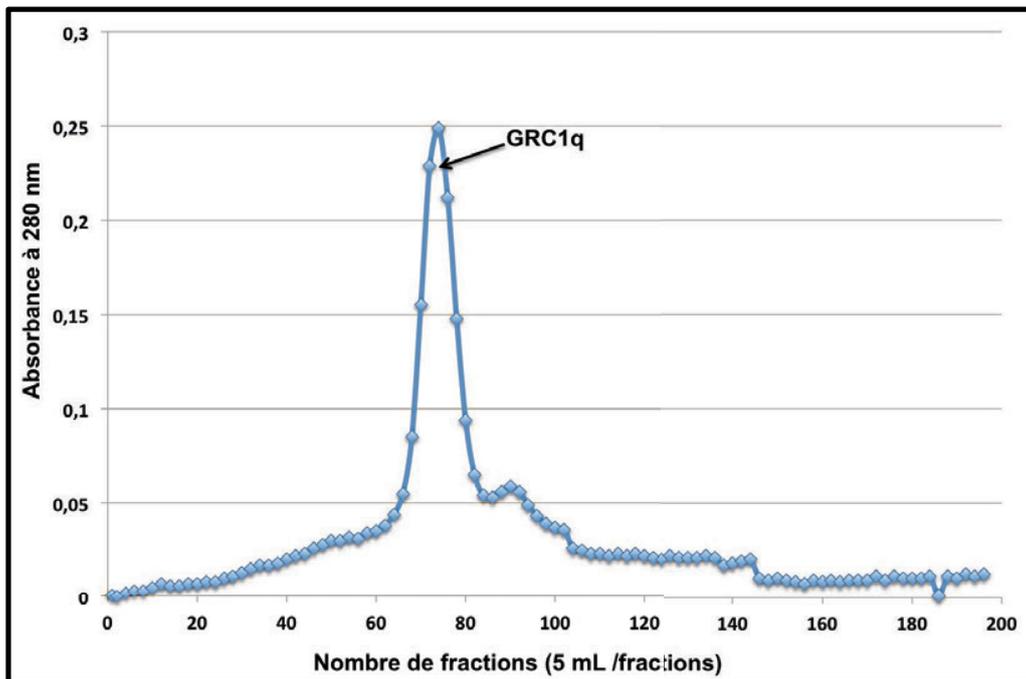


FIGURE 3.VII.1: CHROMATOGRAMME DE L'ETAPE DE PURIFICATION DES GRC1Q RECOMBINANTES

La lecture de l'absorbance à 280 nm en fonction du volume d'élution montre la présence d'un pic majeur correspondant à une concentration en NaCl de 160mM.

Après rassemblement et concentration des fractions contenant les GRC1q recombinantes, nous avons obtenu 12 mg de protéines. L'échantillon concentré a ensuite été analysé sur gel SDS-PAGE en conditions réduites et non réduites (**Figure 3.VII.2**). Le gel confirme la présence de la protéine pure en grande quantité. De même la migration plus lente de l'échantillon réduit confirme la présence de ponts disulfures.

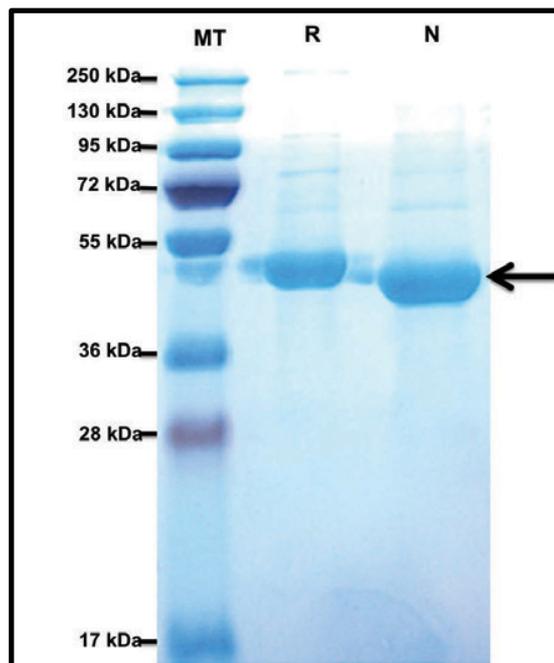


FIGURE 3.VII.2 : GEL BILAN DE LA PURIFICATION DES GRC1Q RECOMBINANTES

Un volume correspondant à 5 µg de protéines a été déposé sur gel SDS PAGE en condition réductrice et non réductrice. Une bande intense migrant à une distance correspondant à la taille attendue pour les GRC1q est présente. (MT : Marqueur de Taille, N : échantillon Non réduit, R : échantillon Réduit)

VIII. Caractérisation des GRC1q recombinantes

Afin de confirmer la nature et l'intégrité de la protéine purifiée, différentes expériences de caractérisation ont été effectuées.

VIII.1. Vérification de la masse et de la séquence N-terminale de la protéine

La séquence N-terminale DQPR déterminée sur la plate-forme de séquençage N-terminal de l'ISBG correspond bien à celle des GRC1q recombinantes. De la même manière, les analyses de spectrométrie de masse MALDI-TOF (réalisées sur la plate-forme de spectrométrie de masse de l'ISBG) ont confirmé la présence d'une espèce unique en solution possédant une masse de 47534 Da pour une masse théorique de la chaîne polypeptidique de 45523,7 Da. Cette différence de 2010 Da correspond vraisemblablement à une glycosylation sur l'asparagine 124 de la chaîne A, similaire à celle présente sur les GRC1q plasmatiques (2059 Da) (Pflieger et al., 2010).

VIII.2. Contrôle qualité RMN 1D

Le bon repliement tridimensionnel des GRC1q recombinantes a été confirmé par une analyse RMN 1D (effectuée sur la plate-forme RMN de l'ISBG) en observant le déplacement chimique des protons ^1H . Dans le cas d'une protéine dépliée, tous les protons sont exposés au solvant et voient un même environnement ; de ce fait les déplacements chimiques correspondant aux protons sont tous regroupés dans la même zone (entre 7,9 et 8,6 ppm). Si la protéine est bien repliée, les protons sont dans des environnements différents (face au solvant, impliqués dans des structures secondaires, etc.). Dans ce contexte, les déplacements chimiques des protons sont répartis sur une vaste plage allant couramment de 6,5 à 10 ppm. Sur l spectre RMN 1D de la protéine (non montré), la dispersion des déplacements chimiques confirme le bon repliement de la protéine.

VIII.3. Etudes fonctionnelles des GRC1q recombinantes

Après avoir vérifié la pureté, la séquence et la présence de modifications post-traductionnelles, ainsi que le bon repliement de la protéine, nous avons testé sa fonctionnalité. Pour cela, nous avons comparé l'interaction des GRC1q recombinantes et des GRC1q plasmatiques avec des ligands connus (HuCRT et le récepteur des GRC1q : gC1qR, PTX et l'héparine) par SPR. Ces expériences nous ont permis de confirmer la capacité d'interaction des GRC1q recombinantes, en comparaison avec les GRC1q plasmatiques, comme le montre la **Figure 3.VIII.1**. Nous avons également pu montrer que les deux formes des GRC1q, recombinantes et plasmatiques interagissaient avec des constantes d'affinités similaires (I. Bally et N. Thielens).

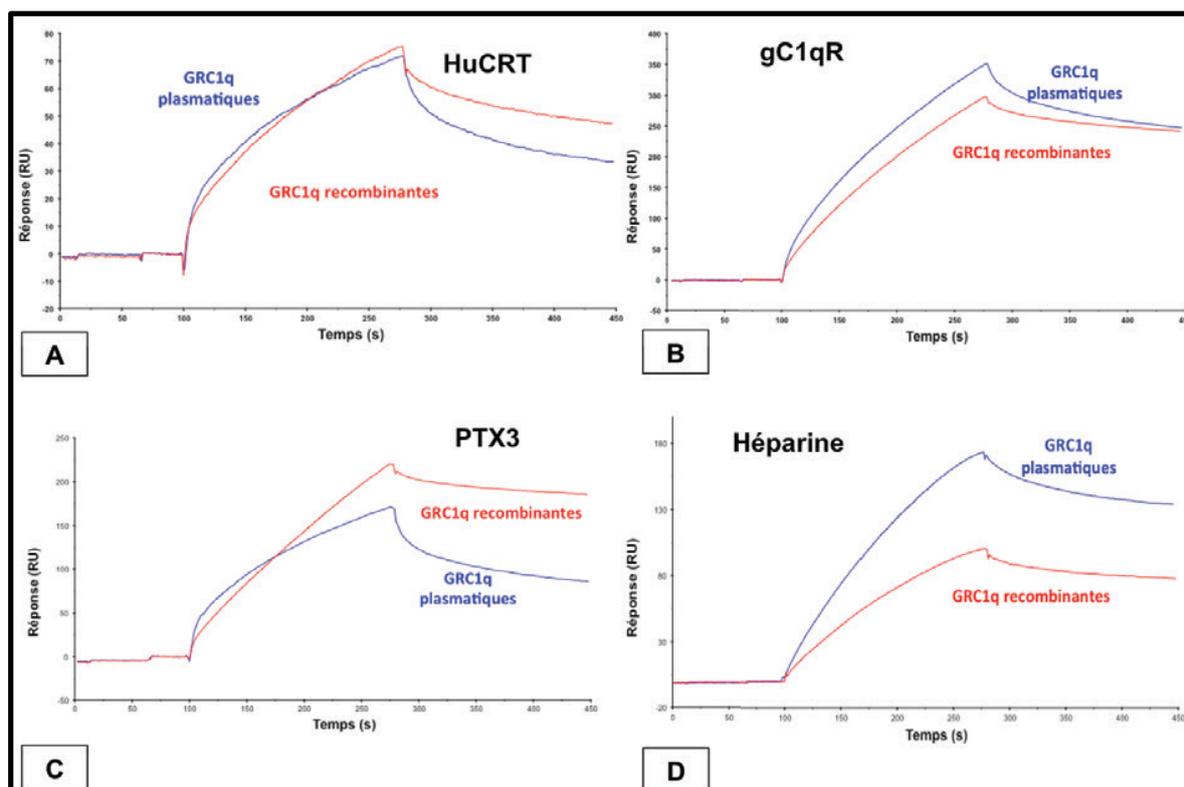


FIGURE 3.VIII.1: CARACTERISATION FONCTIONNELLE DES GRC1q RECOMBINANTES.

Les GRC1q recombinantes (rouge) ou plasmatiques (bleue) ont été injectés en tampon HBS-P sur une puce sur laquelle sont immobilisés la HuCRT (**A**), le récepteur gC1qR (**B**), PTX3 (**C**) et l'héparine (**D**) (concentrations de GRC1q injectées: 1200nM (A et B), 800nM (1600nM pour les plasmatiques) (C) et 240nM (180nM pour les plasmatiques) (D).

VIII.4. Etudes structurales des GRC1q recombinantes

Nous avons ensuite choisi d'étudier la structure de la forme recombinante des GRC1q, afin d'observer par exemple l'effet des linkers qui pourraient créer des contraintes physiques sur la chaîne principale et modifier l'orientation de certains acides aminés. Nous avons donc décidé de résoudre la structure des GRC1q recombinantes à un niveau atomique par cristallographie aux rayons X.

Pour cela les GRC1q recombinantes ont été soumises à un criblage des conditions de cristallisation sur la plateforme HTX de l'EMBL à Grenoble à une concentration de 5 mg/mL et à 20°C. Après quelques jours, de nombreux cristaux sont apparus, que nous avons décidé de reproduire manuellement. Les conditions de cristallisation identifiées ne contenant pas de calcium, nous avons décidé d'en rajouter dans certains essais lors de cette phase de reproduction. De façon intéressante, les cristaux ont été obtenus dans des conditions très

proches de celles ayant permis d'obtenir les cristaux des GRC1q plasmatiques et possédaient souvent la même forme (**Gaboriaud et al., 2003**) (**Figure 3.VIII.2**). Cependant nous n'avons pas enlevé l'acide sialique de la protéine recombinante, contrairement aux GRC1q plasmatiques qui avaient été traitées par la sialidase.



FIGURE 3.VIII.2 : PHOTO DE CRISTAUX DE GRC1Q RECOMBINANTE

Nous avons été en mesure de reproduire les cristaux dont les premiers sont apparus dès 12 h. Différents cristaux ont été testés et nos meilleurs jeux de données, complets à une résolution de 1,55Å et 1,35Å, ont pu être obtenus sur la ligne de lumière ID23-1 de l'ESRF à partir d'un cristal ayant poussé en absence de calcium dans les conditions : **30% PEG 8000, 0,1 M Hepes pH 7 à 20°C** et d'un second cristal ayant poussé en présence de calcium (**23% PEG3350, 0,1 M Tris, 0,2 M NaCl, 50 mM CaCl₂ pH 8,5**). Les statistiques de ces jeux sont présentées dans le **tableau 3.VIII.1**. L'empilement cristallin possède un axe de symétrie d'ordre 2 et s'organise dans un groupe d'espace C2.

La résolution de la structure des GRC1q recombinantes par remplacement moléculaire, dont les statistiques de l'affinement sont présentées dans le **tableau 3.VIII.2**, nous a permis de vérifier que les chaînes principales des GRC1q plasmatiques et des GRC1q recombinantes se superposaient parfaitement (déviations des C α de 0,109 Å sur 331 résidus) (**Figures 3.VIII.4**). De la même manière, de nombreuses chaînes latérales possèdent une orientation et une position identiques et les nombreuses molécules d'eau structurantes sont également présentes aux mêmes positions, à une exception que je détaillerai par la suite. Lors de la résolution de la structure, nous n'avons apporté que de très légères modifications par rapport au modèle de départ issu du remplacement moléculaire. Deux différences ont néanmoins été relevées. La première est la présence des linkers dans la forme recombinante, ces derniers étant relativement flexibles ne possédaient pas toujours une densité électronique bien définie. Nous

| | GRC1q recombinante (condition sans calcium) | GRC1q recombinante (condition avec calcium) |
|---|--|--|
| Ligne de lumière | ID23-1 (ESRF) | ID23-1 (ESRF) |
| Date de la collecte | 15/06/15 | 15/06/15 |
| Longueur d'onde (Å) | 0.976 | 0.976 |
| Détecteur | Pilatus | Pilatus |
| Groupe d'espace | C2 | C2 |
| Paramètres de mailles (a, b, c) (Å) | 81.8, 52.9, 89.9 | 81, 52.7, 89.8 |
| Paramètres de mailles (α , β , γ) (°) | 90, 115.2, 90 | 90, 115.2, 90 |
| Résolution (haute/basse) (Å) | 1.55-100 | 1.35-100 |
| Dernière tranche de résolution (Å) | 1.55-1.61 | 1.35-1.4 |
| Réflexions observées | 243827 (25061) | 363221 (31079) |
| Réflexions uniques | 49348 (5211) | 74618 (7346) |
| Redondance | 4.9 (4.7) | 4.8 (3.7) |
| Complétude (%) | 98.4 (96.7) | 97.7 (85.3) |
| I/σ | 14.85(1.95) | 11.29(1.53) |
| Rsym (%) | 5.6 (75.3) | 7 (83.5) |
| CC (1/2) | 99.9 (56) | 99.8 (68.2) |
| Facteurs B Wilson | 27 | 21 |

TABLEAU 3.VIII.1: STATISTIQUES DES DONNEES DE DIFFRACTION DES JEUX NATIFS DES GRC1Q RECOMBINANTES. LES VALEURS ENTRE PARENTHESES CORRESPONDENT À LA DERNIERE TRANCHE DE RESOLUTION

| GRC1q recombinante | Mg ²⁺ | Ca ²⁺ |
|-----------------------------------|------------------|------------------|
| Résolution (Å) | 1.55 | 1.35 |
| Rfact (%) | 18.01 | 17.95 |
| Rfree (%) | 20.39 | 20.36 |
| Déviations liaisons (Å) | 0.007 | 0.012 |
| Déviations angles (°) | 1.15 | 1.45 |
| Ramachandran outliers (%) | 0.2 (<0,2) | 0 (<0,2) |
| Ramachandran favorisés (%) | 96 (>98) | 95 (>98) |
| Rotamères outliers (%) | 0.8 (1) | 0.8 (1) |
| Clashscore | 4.51 | 2.49 |
| C β outliers | 0 (0) | 0 (0) |
| Facteur B moyen (Å ²) | 27.8 | 21.7 |
| Score Molprobity | 1.55 | 1.39 |

TABLEAU 3.VIII.2: STATISTIQUES D'AFFINEMENT DE LA STRUCTURE DES GRC1Q RECOMBINANTES RESOLUE PAR REMPLACEMENT MOLECULAIRE

avons cependant été en mesure de replacer l'intégralité de la chaîne principale du linker présent entre les chaînes C et B (GSAKA), et les deux derniers acides aminés (KQ) du linker entre les chaînes A et C (GSGKQ) (**Figure 3.VIII.4 A et 3.VIII.5 A**). La présence du linker entre les chaînes C et B (GSGKA) modifie légèrement l'orientation de la chaîne principale en amont et en aval du linker, mais l'orientation de la chaîne latérale des acides aminés est strictement conservée (**Figure 3.VIII.5 B**). Pour le linker entre A et C (GSGKQ), cette fois les chaînes principales en amont et en aval sont très conservées, et une légère différence est observée dans l'orientation de la chaîne latérale de la phénylalanine de la chaîne A (**Figure 3.VIII.5 C**). Néanmoins ces extrémités se poursuivant sous la forme de tiges collagènes dans la molécule de C1q, elles ne constituent pas une surface accessible et il est peu probable qu'elles soient impliquées dans les interactions.

La seconde petite différence visible se situe là-aussi du côté des tiges de type collagène dans le contexte de C1q ; il s'agit de l'extrémité C-terminale de la chaîne B qui, dans le cas des GRC1q recombinantes, est très flexible. Nous n'avons pas été en mesure de reconstruire les deux derniers résidus (M224^B et E225^B) et nous pouvons constater que la chaîne principale est légèrement décalée (**Figure 3.VIII.5 D**).

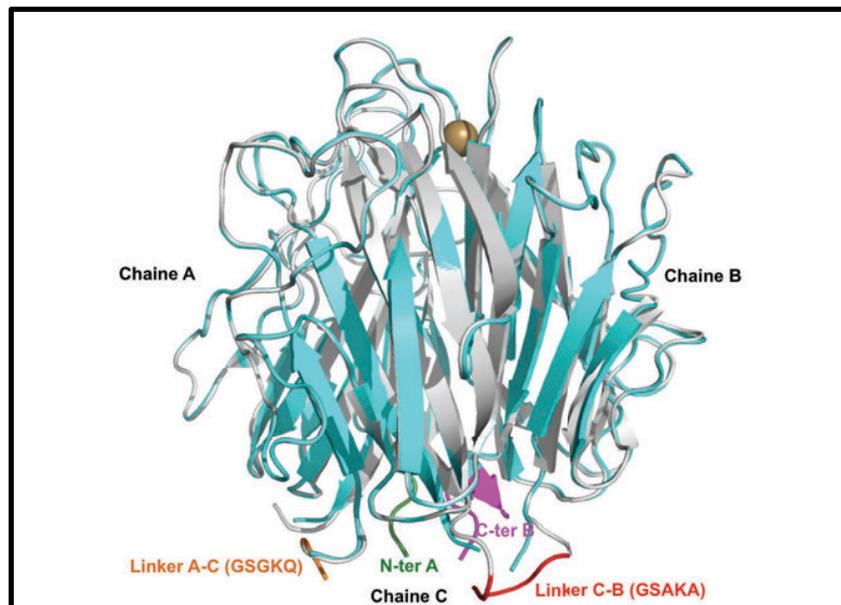


FIGURE 3.VIII.4: SUPERPOSITION DE LA STRUCTURE DES GRC1Q RECOMBINANTES ET PLASMATIQUES

La superposition des structures des GRC1q recombinantes (gris) et plasmatiques (bleu) montre une importante conservation de la chaîne principale et des structures secondaires. Les trois chaînes A-C-B se succèdent dans le sens des aiguilles d'une montre et sont reliées par des linkers (A-C : GSGKQ) et (C-B : GSAKA). Les extrémités N-terminale et C-terminale sont indiquées en vert et magenta respectivement.

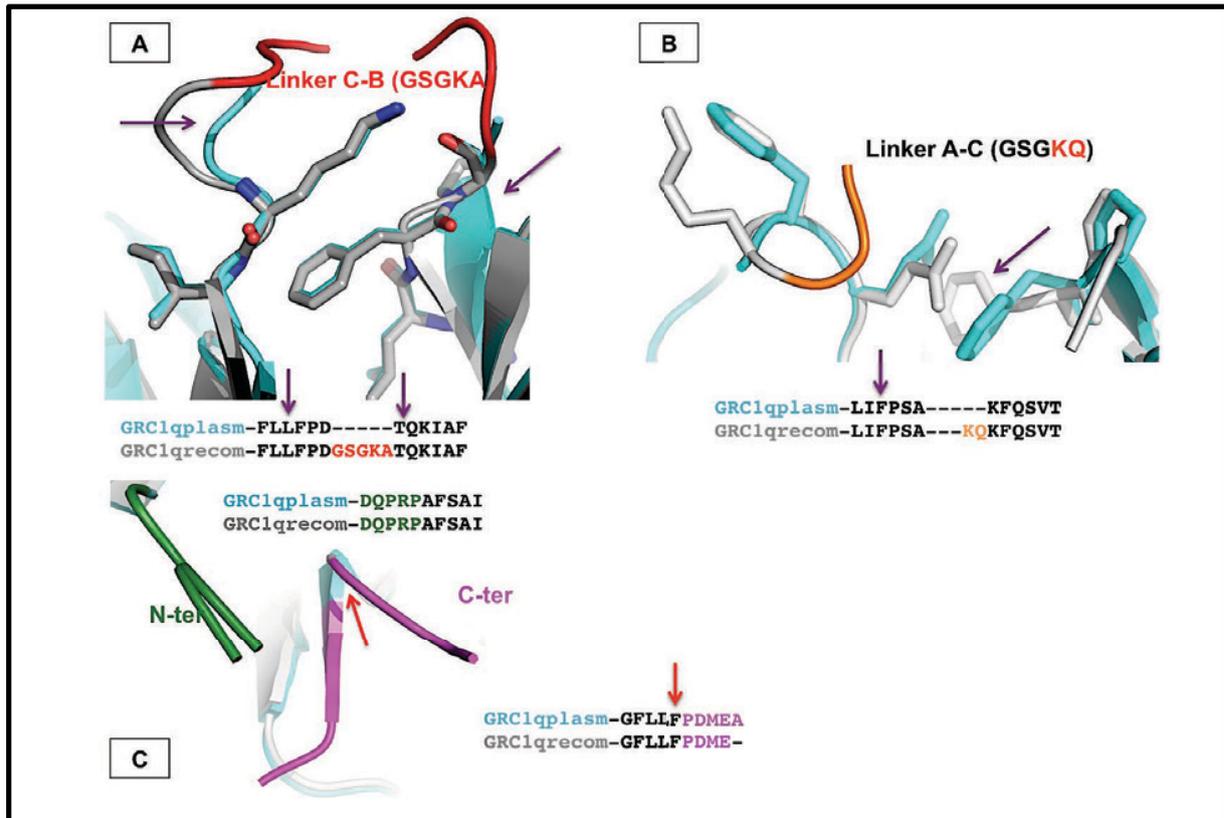


FIGURE 3.VIII.5: DIFFERENCES OBSERVEES DANS LES STRUCTURES DE GRC1Q PLASMATIQUES ET RECOMBINANTES.

La superposition des structures des GRC1q recombinantes (gris) et plasmatiques (bleu) montre certaines différences. (A) La présence du linker entre la chaîne B et C (**GSGKA**) entraîne une déviation de la chaîne principale (boucle de la chaîne C et extrémité du feuillet β pour la chaîne B) (**flèche violette**), néanmoins les chaînes latérales des résidus en amont et aval du linker sont strictement conservées. (B) Le linker entre les chaînes A et C dont les deux derniers résidus ont été reconstruits (**GSGKQ**) est très flexible, mais aucune déviation des chaînes principales en amont et aval n'est observée. Seul le groupement phényle de la phénylalanine de la chaîne C est légèrement décalé (**flèche violette**). (C) L'extrémité **C-terminale de la chaîne B (magenta)**, est très flexible dans les GRC1q recombinantes et montre une déviation de la chaîne principale au niveau d'une phénylalanine entraînant une rupture du petit feuillet β (**flèches rouges**), alors que l'extrémité **N-terminale (vert)** est parfaitement conservée. Les résidus au niveau desquels la chaîne diverge sont représentés par des flèches dans les alignements structuraux.

Ces structures nous ont également permis d'effectuer des comparaisons fines avec celle des GRC1q plasmatiques. Dans la première structure obtenue en absence de calcium, nous avons observé un pic positif dans la carte Fo-Fc correspondant précisément au site de fixation du calcium présent dans les GRC1q plasmatiques (**Figure 3.VIII.6 (A)**). Nous avons spontanément ajouté un ion Ca^{2+} en estimant qu'il pouvait avoir été piégé dans le milieu de culture. Néanmoins le calcul de la nouvelle carte de différence fait apparaître un pic négatif (**Figure 3.VIII.6 (B)**). Nous avons alors, avec l'aide du logiciel Phenix, testé différents ions

et l'ajout d'un ion Mg^{2+} semblait le mieux correspondre à la densité et nous permet d'améliorer les facteurs R de 2% (**Figure 3.VIII.6 (C)**). La présence d'un autre ion divalent ne modifie cependant ni l'orientation des chaînes latérales constituant le site de fixation, ni les chaînes aliphatiques voisines par rapport aux GRC1q plasmatiques (**Figure 3.VIII.6 (D)**). Nous avons été en mesure de ne relever qu'une différence : alors que la quasi-totalité des molécules d'eau est strictement conservée, une molécule d'eau supplémentaire (104) a été identifiée dans la structure des GRC1q recombinantes en présence de magnésium. Sa présence semble modifier légèrement l'orientation de la chaîne latérale du résidu Q173^C (**Figure 3.VIII.7 (D)**).

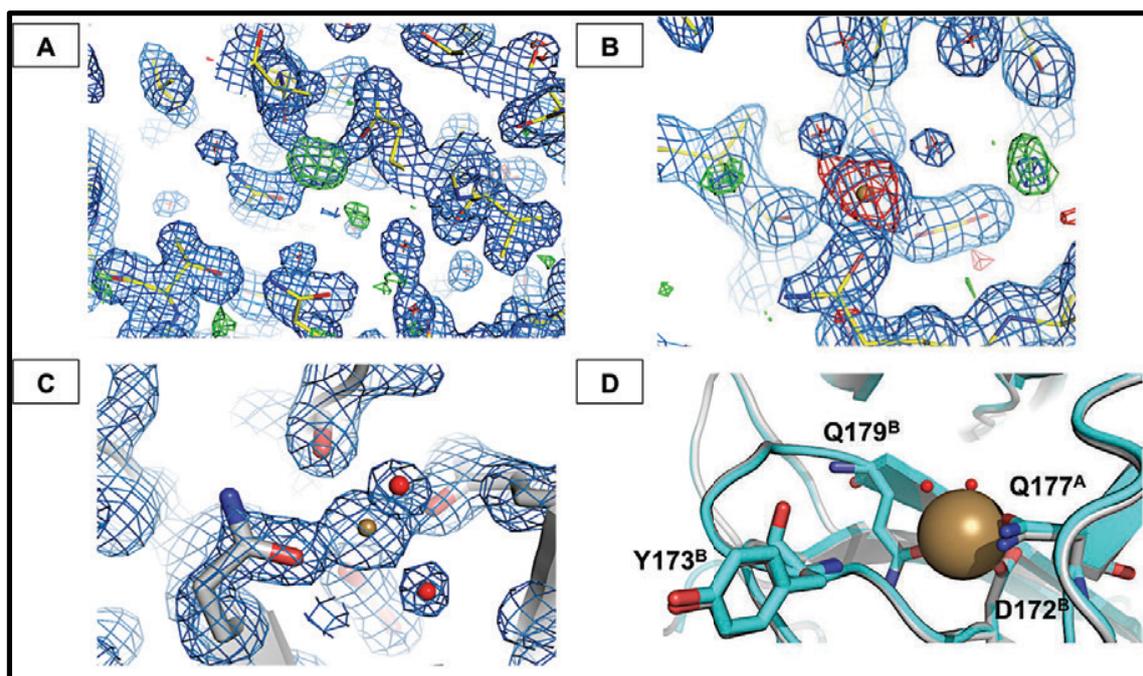


FIGURE 3.VIII.6: SITE DE FIXATION DU CALCIUM DANS LES GRC1Q RECOMBINANTES CRISTALLISEES EN ABSENCE DE CALCIUM

(A) La carte de différence Fo-Fc présente un pic positif au niveau du site de fixation du calcium et en absence de ligand modélisé. (B) Un ion calcium ne s'ajuste pas à la densité électronique, comme cela est caractérisé par le pic négatif de la carte Fo-Fc. (C) Un ion magnésium correspond d'avantage à la densité électronique observée et permet l'amélioration des facteurs R de 2%. (D) La présence d'un ion calcium ou magnésium ne modifie pas la conformation des résidus impliqués dans le site de fixation des GRC1q plasmatiques (bleu) et recombinantes (gris).

La résolution de la structure obtenue en présence de calcium montre que l'ajout d'un ion magnésium entraîne l'apparition d'un pic positif dans la carte de différence Fo-Fc (**Figure 3.VIII.5 (A)**), alors qu'un ion calcium s'ajuste parfaitement à la densité (**Figure 3.VIII.5 (B)**). Cette observation met en évidence la substitution possible de l'ion divalent et nous donne la possibilité de remplacer le magnésium par le calcium physiologique en effectuant

une dialyse. A nouveau, les résidus impliqués dans le site de fixation conservent la même conformation (**Figure 3.VIII.7 (C)**) et la molécule d'eau 104 observée en présence de magnésium est à nouveau présente, tandis que le résidu Q173^C adopte une conformation intermédiaire à celle observée dans les GRC1q plasmatiques et les GRC1q recombinantes en présence de magnésium (**Figure 3.VIII.7 (D)**).

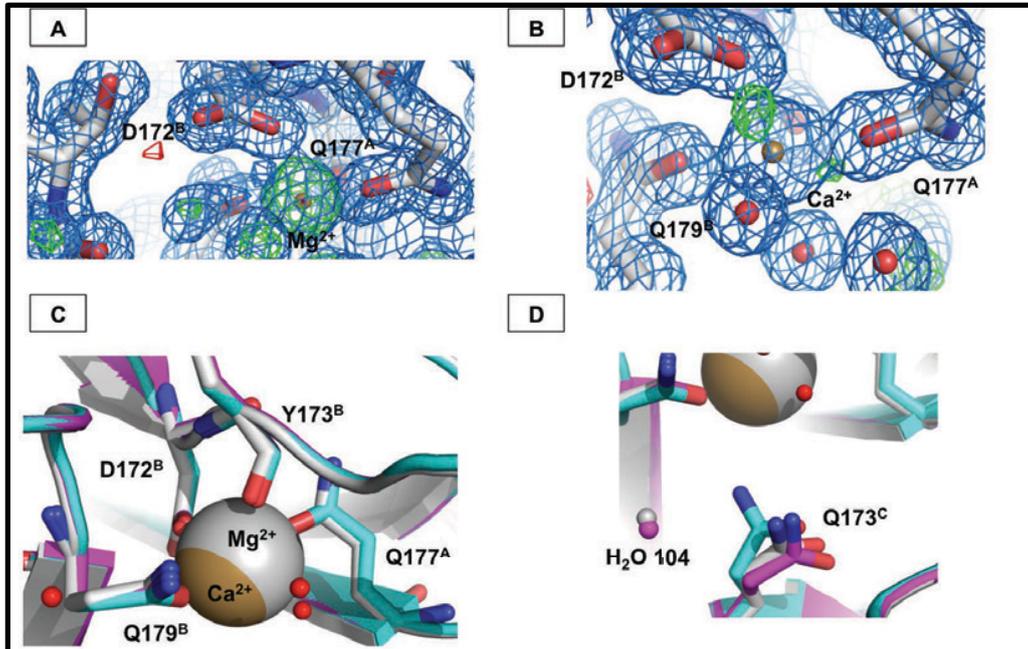


FIGURE 3.VIII.7: SITE DE FIXATION DU CALCIUM DANS LES GRC1Q RECOMBINANTES CRISTALLISEE EN PRESENCE DE CALCIUM

(A) La carte de différence Fo-Fc présente un pic positif au niveau du site de fixation du calcium en présence d'un ion Mg²⁺. (B) L'ion calcium s'ajuste parfaitement à la densité électronique. (C) Les ions Mg²⁺ (gris) et Ca²⁺ (doré), qu'il soient issus des GRC1q plasmatiques ou recombinantes se superposent parfaitement. La conformation des résidus constituant le site de fixation du calcium est extrêmement conservée pour les GRC1q plasmatiques (cyan), les GRC1q recombinantes en présence de Ca²⁺ (magenta) et les GRC1q recombinantes en présence de Mg²⁺ (gris). (D) La molécule d'eau 104 est conservée dans les GRC1q recombinantes uniquement, qu'elles soient en présence (magenta) ou en absence (gris) de Ca²⁺. La présence de la molécule d'eau entraîne légère modification de l'orientation de la chaîne latérale du résidu Q173 de la chaîne C.

4. Discussion

Nous allons ici conclure sur les différents objectifs proposés en fin d'introduction pour ce travail de thèse débuté en octobre 2012. Nous discuterons ensuite plus particulièrement des résultats obtenus par analyse structurale des calréticulines et sur les hypothèses que l'on peut proposer quant aux interactions entre les différentes calréticulines étudiées et C1q. Pour finir, nous évoquerons les perspectives d'analyses fonctionnelles avec nos collaborateurs.

Pour ce qui est de l'analyse des domaines impliqués dans l'interaction entre C1q et les calréticulines de parasite, nous avons pu montrer que les domaines globulaires avaient un rôle important, ce qui nous a incité à approfondir l'étude structurale des domaines globulaires de calréticuline et développer une forme recombinante monocaténaire des domaines globulaires de C1q.

I. Forme monocaténaire des régions globulaires de C1q

Nous avons vu à la fin de la section Résultats qu'il a été possible de produire en grande quantité cette forme recombinante dans le système d'expression HEK-293F (10 mg à partir de 500 mL de surnageant de milieu de culture). Nous avons pu contrôler la masse, la séquence, la présence de glycosylation et même la structure à une résolution atomique de notre construction, le tout correspondant parfaitement aux GRC1q plasmatiques à l'exception de la présence des linkers. De manière encourageante les premiers tests fonctionnels avec cette forme recombinante ont permis de confirmer ses propriétés d'interaction par comparaison avec les GRC1q purifiées à partir du plasma. La substitution de l'ion calcium par un autre ion divalent ne modifie ni la conformation des acides aminés impliqués dans la fixation, ni la structure tertiaire au voisinage. De même, la présence des linkers ne semble pas modifier leur environnement direct.

II. Etudes des calréticulines de *T. cruzi* et *E. histolytica*

I.1. Analyses structurales

Un des objectifs de départ qui me tenait particulièrement à cœur était de résoudre la structure de la calréticuline de *T. cruzi* à un niveau moléculaire grâce à la cristallographie aux rayons X. En me basant sur les stratégies employées pour obtenir les structures des calréticulines

humaine et de souris, j'ai procédé à une dissection des domaines flexibles de la TcCRT pour obtenir un fragment capable de cristalliser. Il s'est avéré que des troncations importantes de l'hélice C-terminale provoquaient la déstabilisation de la protéine. J'ai pu néanmoins après un an et demi d'échecs obtenir des premiers cristaux grâce à des études complémentaires de la flexibilité de la protéine, qui m'ont incité à réduire le domaine P. Une fois la structure d'un fragment de TcCRT obtenue, j'ai pu cristalliser rapidement le fragment homologue de la EhCRT, dont j'avais réalisé la construction en parallèle. Nous avons ainsi obtenu les structures de ces deux fragments incluant le domaine globulaire, la jonction et un petit fragment de P, ce qui est utile pour analyser les positions relatives des domaines. La comparaison des différentes structures de CRTs nous a permis de constater la stabilité de la jonction et son rôle dans le maintien de l'orientation relative du domaine P par rapport au domaine globulaire. Cette jonction présente des conformations légèrement plus variées dans le cas de la EhCRT, probablement à cause de la présence de l'alanine 197 (**Figures 3.III.21**). Une différence claire existe à ce niveau entre les calréticulines et les protéines homologues de la famille de la calnexine. Cette divergence ne s'explique pas par la séquence primaire de la partie C-terminale de la jonction, celle-ci étant plus proche entre la CNX et les CRTs de parasite qu'entre les CRT de parasite et les CRT de mammifère. Néanmoins 2 motifs uniques à la CNX semblent induire l'orientation du domaine P (**Annexes 5 et 6 pages E à P**). La première est la présence d'une hélice précédant la partie N-terminale de la jonction dans les calréticulines. Comme indiqué dans la section Résultats, cette hélice est constituée de résidus acides formant un cluster avec lequel interagit le premier résidu basique de la jonction pour la maintenir vers le domaine globulaire. Dans la calnexine, cette hélice n'a pas été définie dans la structure, cependant l'absence des résidus acides, substitués par des résidus hydrophobes, et la présence d'une insertion pourraient expliquer l'éloignement de la jonction et du domaine P par rapport au domaine globulaire (**Figure 4.II.1 (A)**). La CNX possède également une insertion de 6 résidus formant une boucle supplémentaire semblant influencer l'orientation du domaine P en exerçant une attraction vers ce domaine (**Figure 4.II.1 (B)**).

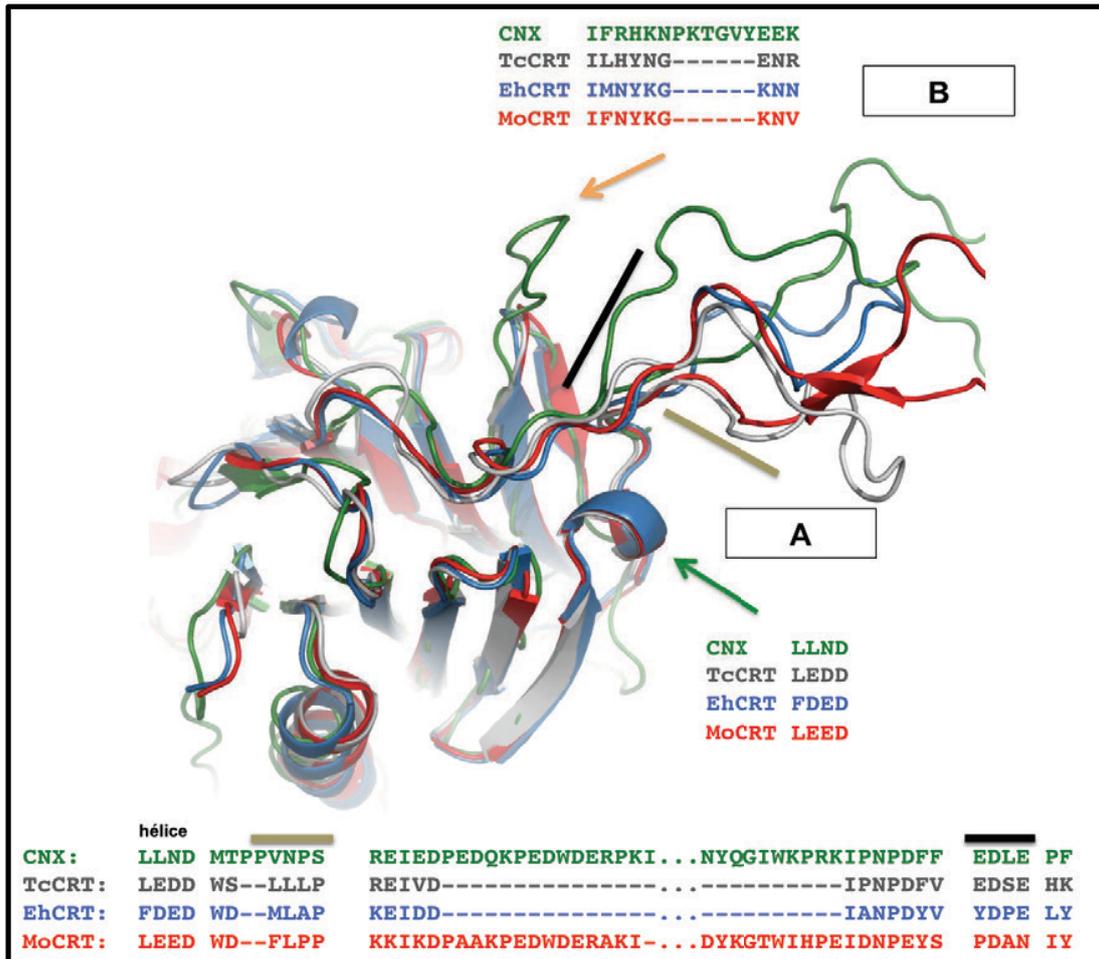


FIGURE 4.II.1 : DIFFERENCE DANS L'ORIENTATION DE LA JONCTION ENTRE LA CNX ET LES CRTs

La position de la jonction dans la **CNX (vert)** est légèrement décalée et modifie l'orientation du domaine P qui n'est pas ramené vers le domaine globulaire contrairement aux CRTs. La conformation de la jonction de la **EhCRT (bleu)** se rapproche de celle de la **MoCRT (rouge)** ou de la **TcCRT (gris)** même si elle est légèrement décalée, probablement à cause de l'alanine présente dans la partie **N-terminale** de la jonction (**barre dorée**). La séquence de la partie **C-terminale** de la jonction est conservée entre la CNX et les CRTs de parasite (**barre noire**). **(A)** L'hélice précédant la partie N-terminale de la jonction crée un cluster de résidus acides maintenant le domaine P vers le domaine globulaire. Ces résidus acides sont absents de la CNX. **(B)** Une insertion dans la CNX crée une boucle supplémentaire contraignant l'orientation du domaine P.

La résolution de la structure des deux CRTs de parasite nous a permis de mettre en évidence de légères différences entre elles et avec les protéines de mammifères. Ces divergences correspondent à des insertions et des délétions situées exclusivement dans des boucles et sur la même face. Les calréticulines étant des protéines multifonctionnelles, ces régions pourraient être potentiellement impliquées dans des fonctions n'étant pas directement associées à C1q, fixé par les 3 CRTs.

Les différentes conformations observées pour les monomères de EhCRT, couplées aux études de SAXS effectuées sur différents fragments de la TcCRT, nous ont permis de percevoir des éléments de la dynamique du domaine P, pour lequel une charnière flexible est située en aval de la jonction conservée et rigide. Autour de cette charnière, le domaine P s'ouvre et se replie comme un levier. En absence de substrat, le domaine P oscille entre des positions totalement ouvertes, permettant à la calréticuline d'accueillir un substrat, et partiellement fermées, réduisant la possibilité de fixation. Cela laisse supposer que la protéine doit être associée à un partenaire au niveau du domaine P pour le maintenir dans sa conformation ouverte. Ce partenaire pourrait être la protéine co-chaperonne ERp57, qui s'associe à l'extrémité du domaine P (Wijeyesakere et al., 2013).

La conformation fermée de la EhCRT nous a donné une idée sur le mécanisme de fermeture du domaine P sur un substrat. Cette conformation a introduit le seul cas de figure dans lequel la jonction du domaine P se réorganise. Ici, la chaîne principale des acides aminés constituant la jonction se déplace d'environ 4 Å et entraîne la rupture de la petite hélice $\alpha 1$, absente de la CNX (**Figure 3.III.23 (B)**). Il se pourrait que la présence d'un substrat entraîne une modification de l'environnement des acides aminés de la jonction, provoquant un réarrangement conformationnel de cette dernière et aboutissant au basculement du domaine P le long de la charnière. Néanmoins cette observation a été effectuée sur un fragment possédant un domaine P tronqué et il est possible que le repliement du domaine P complet n'entraîne pas des contraintes aussi importantes sur la jonction.

Nous nous sommes également intéressés à une autre région flexible : l'hélice C-terminale. L'alignement des structures obtenues pour les différentes CRT dans le cristal désigne une charnière de flexibilité à l'endroit où l'hélice n'est plus stabilisée par le domaine globulaire. La position de cette charnière est confirmée dans les études SAXS sur la protéine entière dans lesquelles le modèle s'ajuste aux données expérimentales par un déplacement de l'hélice au niveau d'une charnière équivalente. Étonnamment, pour la TcCRT cristallisée, la courbe de diffusion du modèle s'ajuste en proposant une charnière située plus bas. Il se pourrait que les coupures de l'hélice déstabilisent son interaction avec le domaine globulaire et lui permette de s'écarter de ce dernier. Il faut toutefois rappeler qu'à la résolution du SAXS un décalage de quelques acides aminés pour la charnière n'est peut-être pas significatif.

L'hélice C-terminale serait donc flexible en absence de molécules stabilisantes à son voisinage. Cette extrémité C-terminale a été décrite comme étant potentiellement responsable de la retranslocation cytoplasmique de la calréticuline en dehors du RE. Cette réponse serait induite par un stress lié à une diminution drastique du calcium. Les ions Ca^{2+} au voisinage de

l'hélice la stabilisent (les sites de fixation du calcium C-terminaux ne sont pas présents dans le fragment cristallisé). La diminution de cette stabilisation semble augmenter sa flexibilité et, de manière concomitante, le dépliement partiel de l'hélice C-terminale de la CRT, ce qui provoque son export via la machinerie ERAD (**Labriola et al., 2010**).

II.2. Comparaison des structures en lien avec la fonction chaperonne

En contraste avec la flexibilité des domaines P et C-terminaux, nous avons pu confirmer la conservation du domaine lectine et du site de fixation spécifique du substrat Glc₁Man₄ chez deux espèces phylogénétiquement très éloignées des mammifères. Nous avons d'ailleurs vu en introduction des différences importantes entre ces espèces et les mammifères, comme par exemple l'utilisation de la néoglycosylation et l'absence de calnexine dans le système. Cette importante conservation structurale du domaine lectine apparaît donc d'autant plus remarquable.

Nous pouvons confirmer la position homologue du site de fixation du glucose, même si nos structures montrent des variations inattendues dans le mode précis de reconnaissance de ce sucre (**Figure 3.III.10**). Nous avons également observé la fixation de l'ion chlorure à proximité immédiate de ce site, lequel pourrait participer au réseau de liaisons hydrogène. Enfin, parmi les observations imprévues réalisées pendant cette thèse, nous pouvons rappeler que la molécule identifiée d'après la densité comme du glucose dans une des structures de EhCRT a probablement été fortement fixée et transportée depuis les étapes de production de la protéine jusqu'à celles de la cristallisation.

A côté de ce mécanisme dépendant de la présence du glucose, il est aujourd'hui admis que la calréticuline joue son rôle de protéine chaperonne pour des protéines non glycosylées à partir d'un second site de reconnaissance. La localisation de ce site a longtemps été sujette à discussion et de récents travaux ont pu proposer et confirmer sa position à proximité du site lectine (**Chouquet et al., 2011; Wijeyesakere et al., 2013**). Lors de la résolution des structures de la EhCRT et de la TcCRT, nous avons observé de façon répétée des interactions entre le site lectine et des fragments de domaine P des molécules voisines exposant des chaînes latérales hydrophobes. Ceci confirme que les CRTs de parasite sont capables de fixer des peptides similaires à des protéines partiellement dénaturées sur le bord du domaine lectine et à proximité du site de fixation du glucose. Ces différentes observations dans les empilements cristallins de MoCRT, HuCRT (**Chouquet et al., 2011**), EhCRT et TcCRT montrent l'implication commune de certains résidus, qui corrélerent avec des données fonctionnelles *in vivo* de la fonction chaperonne (W314 et le pont disulfure) (**Martin et al.,**

2006). Ce sujet d'étude étant assez éloigné du centre d'intérêt actuel de l'équipe, nous avons établi des contacts pour réaliser des expériences complémentaires dans le cadre d'une collaboration afin de confirmer si la substitution des acides aminés identifiés affecte le repliement de protéines non glycosylées. Les expériences de mutation pourraient également inclure les acides aminés qui nous paraissent importants pour la dynamique du domaine P (Wijeyesakere et al., 2013).

III. Analyses d'interaction *in vitro* avec C1q

L'identification des domaines de la calréticuline et de C1q impliqués dans l'interaction de ces deux protéines a été sujet à de nombreuses controverses au cours des quinze dernières années. En effet, il a d'abord été proposé que la reconnaissance s'effectuait essentiellement soit par les domaines collagènes de C1q (Malhotra et al., 1990; Stuart et al., 1996, 1997), soit par les domaines globulaires de C1q (Kishore et al., 1997; Kovacs et al., 1998). Au final, il a été récemment admis que les GRC1q et les CLR participaient à l'interaction avec la HuCRT (Païdassi et al., 2008a; Steinø et al., 2004). Il a été alors proposé que les GRC1q assuraient la fonction collectine (reconnaissance des cellules apoptotiques par exemple), alors que les CLR interviendraient dans la transduction du signal (interaction avec les récepteurs des macrophages par exemple), comme illustré sur la **Figure 4.III.1**. Du côté de la HuCRT, s'il a longtemps été proposé une implication des domaines globulaires et P (Malhotra et al., 1993; McGreal and Gasque, 2002; Stuart et al., 1996, 1997), il semblerait que cette interaction s'effectue essentiellement à travers le domaine globulaire respectif (Eggleton et al., 1994; Païdassi et al., 2011) et que le domaine P seul n'interagisse pas avec C1q. Nous nous sommes posé la même question avec les deux calréticulines de parasite étudiées. Nous espérons que cela nous donnerait éventuellement des indications plus précises sur les domaines impliqués dans l'interaction entre C1q et les calréticulines de surface fonctionnant comme des signaux "eat-me" (TcCRT et cellules apoptotiques) et dans l'interaction entre C1q et la calréticuline présente à la surface des phagocytes (EhCRT).

Les expériences ELISA et de SPR ont montré et confirmé d'un point de vue qualitatif que les domaines globulaires des CRTs de parasite et les régions globulaires de C1q étaient principalement impliquées dans l'interaction. Ces expériences d'interactions effectuées lors de ma thèse et publiées (Païdassi et al., 2011) ont montré que la HuCRT semble sensible à la délétion de son domaine P dans l'interaction avec C1q contrairement aux calréticulines de parasites. De plus, la HuCRT semble interagir plus fortement avec les CLR en test ELISA.

Ces observations suggèrent un modèle dans lequel la HuCRT pourrait être reconnue par les GRC1q et les CLR. Dans ce cas, C1q pourrait jouer le rôle de molécule pontante à la surface des cellules apoptotiques et des macrophages comme décrit dans la **figure 4.III.1**. Cette interaction impliquerait donc différents modes de la reconnaissance de la CRT.

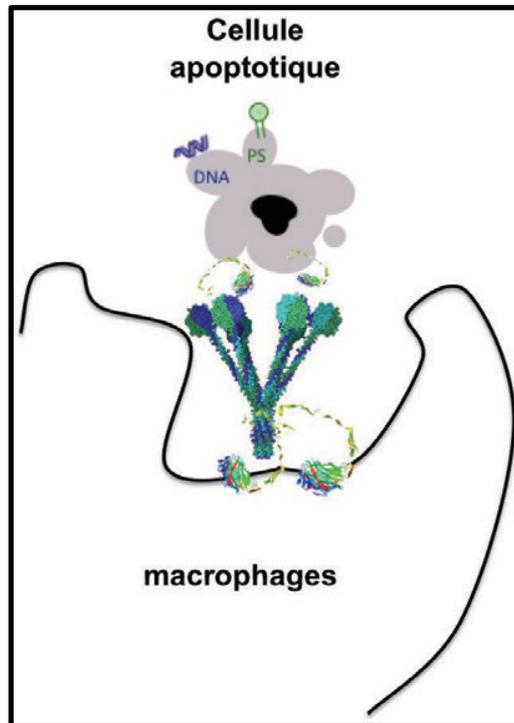


FIGURE 4.III.1 : PREMIER MODELE PROPOSE POUR L'INTERACTION ENTRE LES C1Q ET LA CRT A L'INTERFACE ENTRE LES MACROPHAGES ET LES CELLULES APOPTOTIQUES

Dans ce modèle, incompatible avec les données expérimentales, C1q joue le rôle de molécule pontante entre les cellules apoptotiques et les macrophages, par l'intermédiaire d'une interaction entre la CRT à la surface des deux types cellulaires. Cette interaction implique le domaine globulaire des deux protéines d'une part et les CLR avec le domaine P d'autre part

Toutefois, dans ce modèle, C1q aurait la possibilité de fixer des CRT présentes à la surface avec ses 6 régions globulaires. Or nous avons pu observer lors des expériences de SPR que les affinités étaient similaires entre C1q et les GRC1q. Après avoir longuement discuté avec des membres du laboratoire étudiant les interactions de C1q depuis des années, j'ai pu apprendre que le phénomène d'avidité provoqué par l'hexavalence de la molécule conduit généralement à une augmentation de l'affinité d'un facteur de l'ordre de 100 par rapport aux GRC1q. De plus nous avons pu constater qu'un modèle divalent n'expliquait pas mieux l'interaction qu'un modèle 1 :1. Il est donc possible qu'une seule GRC1q par molécule de C1q soit impliquée dans l'interaction. Les constantes d'affinité similaires entre les espèces nous ont également fait penser à un site de fixation commun impliquant essentiellement le domaine globulaire des CRT. En effet les EhCRT et TcCRT ne semblent pas sensibles à la troncation

du domaine P et interagissent faiblement avec les CLR. Si les deux CRT de parasite miment réellement les différentes fonctions de la protéine humaine, alors les régions collagènes ne devraient pas intervenir, ce qui nous conduit à proposer le modèle d'interaction présenté **figure 4.III.8**, dans lequel la molécule C1q établit un pont entre les cellules apoptotiques et les macrophages en se couchant. Ces résultats sont confortés par les travaux récents sur la calréticuline de *Brugia malayi* (possédant environ 40% d'identité avec les HuCRT, TcCRT et EhCRT, **Annexe 1, page A**) pour laquelle il a été montré que seules les GRC1q intervenaient dans l'interaction (**Yadav et al., 2014**).

Les analyses d'interaction par SPR ne sont possibles que dans une orientation (**Païdassi et al., 2011**), et elles suggèrent l'existence d'un changement conformationnel lors de cette interaction. Les études préliminaires de RMN, réalisées en solution, semblent également suggérer une dynamique lente d'échanges dans le cadre de cette interaction, avec possibilité d'un changement local de conformation, en particulier aux alentours d'un tryptophane (**Figure 3.IV.11**). La dissociation lente du complexe rend difficile l'identification des résidus impliqués dans l'interaction, mais les analyses RMN seront poursuivies. L'attribution des résidus dans le contexte du complexe étant impossible à réaliser du fait de l'instabilité des GRC1q à forte concentration, il faudra probablement réaliser une expérience de titration inverse, pour laquelle l'ajout d'un excès de protéine non marquée pourrait potentiellement faire disparaître les déplacements chimiques des acides aminés fortement impliqués dans le complexe. Nous pourrions ensuite identifier ces résidus par déduction après attribution de la forme libre de la TcCRT.

Des premières indications plus précises proviennent de la mutation de seulement deux charges dans les domaines globulaires de la molécule C1q recombinante, qui abolissent totalement l'interaction avec les domaines globulaires des CRT de parasite. Ceci suggère en particulier une forte composante de type électrostatique dans l'interaction. Nous pouvons donc observer les surfaces électrostatiques des différentes CRT pour identifier des régions exposant des résidus acides et relativement conservés entre les espèces et comportant un tryptophane à proximité (d'après les résultats RMN), qui pourraient éventuellement correspondre au site de fixation de C1q. Les affinités similaires mesurées en SPR pour tous les fragments suggèrent un site de reconnaissance identique et nous orientent vers le domaine lectine hautement conservé. Néanmoins, les différents tests ELISA répétés de nombreuses fois ont montré une interaction plus faible pour la TcCRT (potentiellement liée à une dissociation plus importante): nous chercherons donc à identifier, au sein de ces régions acides, des différences de séquences pour la TcCRT.

L'observation des surfaces électrostatiques des trois protéines vues du dessus (**Figure 4.III.2**) fait apparaître une région acide pour les TcCRT et HuCRT, et plus variable dans la EhCRT, du fait des différentes conformations adoptées par le domaine P (**Figure 4.II.3**). De plus, cette région acide est marquée par la présence d'une lysine pour la TcCRT au sein d'une séquence conservée alors que les deux autres CRT possèdent un acide glutamique (**Figure 4.III.2**). C1q pourrait alors interagir au niveau de ce site par des liaisons ioniques entre les résidus acides conservés (E/D 125 pour les Hu/TcCRT et E 119 pour la EhCRT). Cette hypothèse est confortée par l'observation des empilements cristallins des HuCRT et EhCRT (**Figure 4.III.3**) : en superposant les domaines lectines, nous pouvons observer que le résidu acide conservé au début de la séquence proposée est en interaction avec deux lysines appartenant au monomère voisin. Cela suggère que les acides aminés basiques des GRC1q pourraient interagir ici (**Figure 4.III.4**).

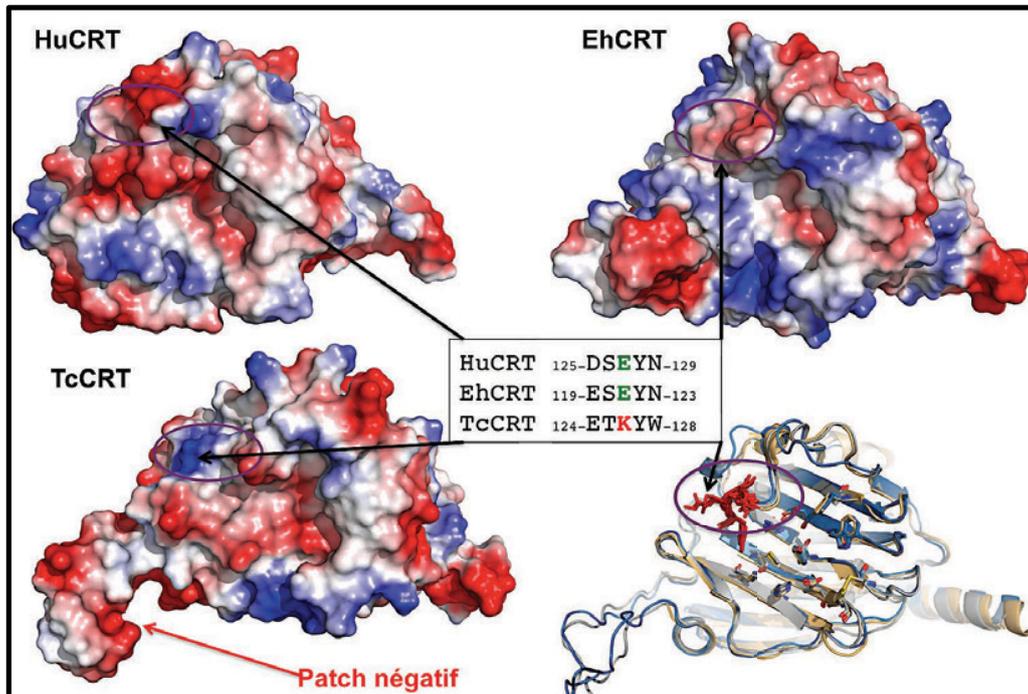


FIGURE 4.III.2 : SURFACE ELECTROSTATIQUE DES CALRETICULINES (VUE DE DESSUS)

Les calréticulines possèdent une région riche en résidus acides (**cercle violet**). Au sein de cette région, la TcCRT possède un acide aminé basique à la place d'un acide contenu dans une séquence conservée et se caractérisant par l'apparition d'un patch positif prononcé. En parallèle, l'intérieur du domaine P, visible dans la structure de la TcCRT et de la EhCRT (chaîne A), présente une surface acide pouvant stabiliser le complexe en se refermant sur les résidus positifs des GRC1q. La séquence identifiée se situe à proximité du domaine lectine.

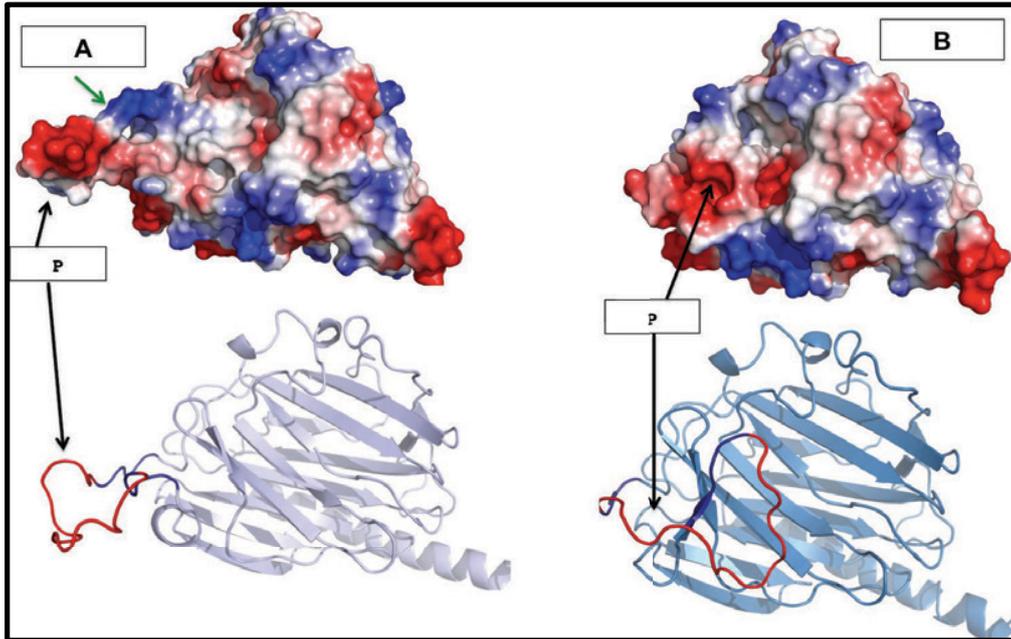


FIGURE 4.III.3 : INFLUENCE DE LA FLEXIBILITE DU DOMAINE P SUR LA SURFACE ELECTROSTATIQUE DE LA EHCRT

Le calcul de la surface électrostatique de la chaîne B de la EhCRT possédant un domaine P ouvert (A) montre un cluster de résidus basiques (flèche verte) sur le bord du domaine lectine. (B) La surface électrostatique de la chaîne C de la EhCRT montre que la fermeture du domaine P masque le cluster positif pour laisser apparaître une surface acide.

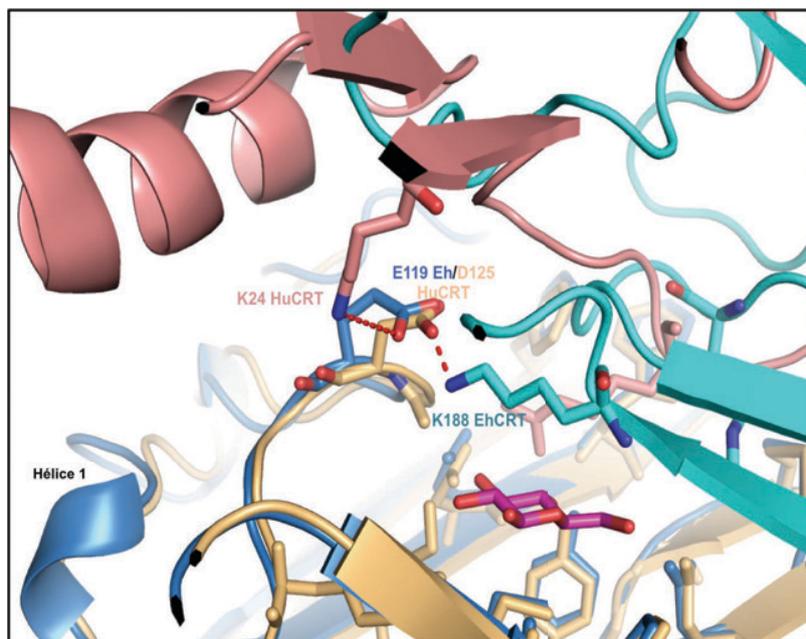


FIGURE 4.III.4 : SITE D'INTERACTION POTENTIEL DE C1Q DANS LA EHCRT ET LA HUCRT

L'observation des interactions dans l'empilement cristallin de EhCRT (bleu/cyan) et de HuCRT (ocre/rosé) montre la participation d'un résidu acide (E119 et D125) proposé comme site d'interaction pour les GRC1q et d'un résidu basique (K24 et K188) pouvant mimer les résidus arginine et lysine des GRC1q.

Nous pouvons également voir sur la **figure 4.III.2** que l'intérieur du domaine P, très visible pour la TcCRT, présente une surface chargée très négativement (**patch acide**). En se référant au modèle dans lequel le domaine P participerait aux interactions avec les protéines non glucosylées en se refermant (**Wijeyesakere et al., 2013**), il se pourrait que le domaine P participe au complexe grâce à la formation de liaisons ioniques avec les résidus basiques des GRC1q. Le résidu acide conservé a néanmoins été muté dans HuCRT, sans effet visible sur l'interaction avec C1q (**annexe 1**). Les CLR semblant contribuer à l'interaction pour la HuCRT, il se pourrait que la mutation d'un site ne soit pas suffisante et cette substitution pourrait avoir d'autres effets sur les CRT de parasite.

Il a été proposé que la HuCRT et le fragment Fc des IgG pourraient partager un site de fixation sur les GRC1q (**Kovacs et al., 1998**). Cette hypothèse semble être confirmée par nos expériences de mutation de C1q inhibant l'interaction avec les CRTs et les IgG/IgM, ainsi que par des expériences de compétition dans lesquelles la pré-incubation de C1q avec les IgG diminue l'interaction avec la CRT de *Brugia malayi* (**Yadav et al., 2014**). Le résidu central de notre séquence présentée **figure 4.III.2** est notamment inclus dans un des sites de fixation de C1q proposés par Yadav et ses collaborateurs (**annexes 1**).

De manière plus intéressante, en regroupant les informations obtenues par RMN et sur les surfaces électrostatiques, nous nous sommes tournés vers une des séquences de fixation de C1q proposée par le groupe de Paul Eggleton (**Kovacs et al., 1998**). Cette séquence contient à nouveau un résidu acide central, excepté pour la TcCRT, où il est remplacé par un résidu basique (**Figure 4.III.5**) et un tryptophane à proximité (193/198/200) (**Figure 4.III.6**). La TcCRT contient en plus un second résidu tryptophane identifié en rouge sur la figure 4.III.6, et même si nous ne pouvons exclure qu'il soit à l'origine du signal observé en RMN, nous discuterons uniquement du W198/193/200, car étant présent dans toutes les CRT et pour les raisons que je vais aborder.

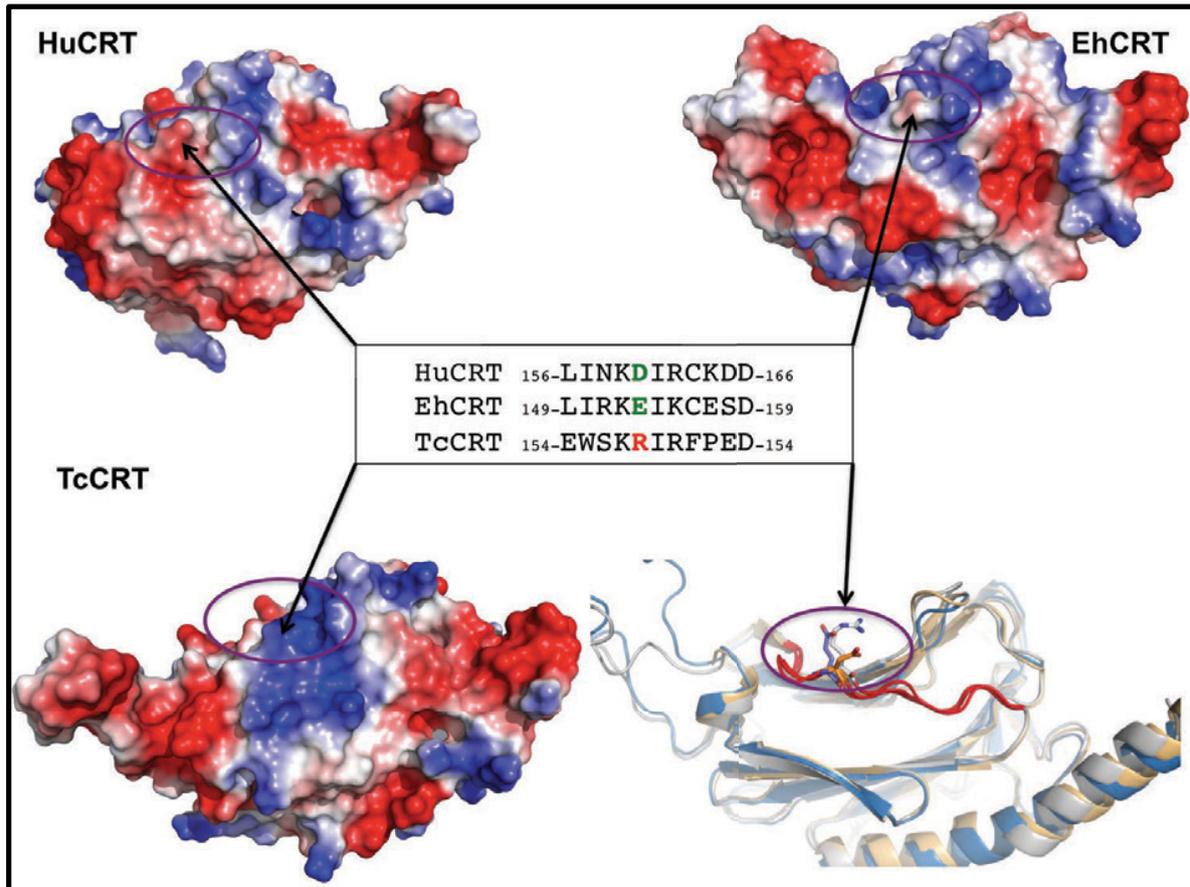


FIGURE 4.III.5: SURFACE ELECTROSTATIQUE DES CALRÉTICULINES (VUE LATÉRALE, DOMAINE P À GAUCHE)

La face latérale des calréticulines (avec le domaine P orienté à gauche) est extrêmement acide. Elle est interrompue au centre par un patch de résidus basiques, très prononcé dans la TcCRT. De façon intéressante, la région acide située entre le domaine P et la surface riche en acides aminés positifs contient une séquence décrite comme un des sites potentiels de fixation de C1q. Au sein de cette séquence, un résidu central acide est substitué en résidu basique dans la TcCRT uniquement (**cercle violet**).

Le résidu tryptophane 198/193 (ou 200 pour la HuCRT), présenté **figure 4.III.6**, se trouve à proximité de l'hélice 1 précédant la jonction. Nous nous sommes alors demandé si toutes nos observations pouvaient être reliées. La fixation de C1q au niveau du site présenté dans la **figure 4.III.5** provoquerait la fermeture du domaine P passant par un réarrangement de l'hélice et de la jonction et une exposition du résidu tryptophane, comme observé en RMN. En effet, nous avons pu observer que dans la chaîne C de la EhCRT, où le domaine P est refermé, la réorganisation de la charnière modifie l'environnement du résidu tryptophane 193 (**Figure 4.III.7 (A)**). En plus la partie N-terminale de la « séquence Eggleton » (LIRK dans la EhCRT et EWSK dans la TcCRT) bascule de 180° dans la chaîne C de la EhCRT (**Figure 4.III.7 (B)**). Ce basculement pourrait être freiné par la présence de la molécule de glucose à

proximité, expliquant potentiellement son effet inhibiteur (**Figure 4.III.7 (B)**). De la même manière un peptide correspondant à la séquence de l'hélice 1 suivie du tryptophane 200 et du début de la jonction pour la HuCRT a été utilisé pour observer l'interaction avec la protéine GABARAP par cristallographie aux rayons X. Ici, le tryptophane est impliqué dans l'interaction, alors qu'il est normalement enfoui dans la protéine native (**Thielmann et al., 2009**) (**Figure 4.III.7 (B) et (C)**).

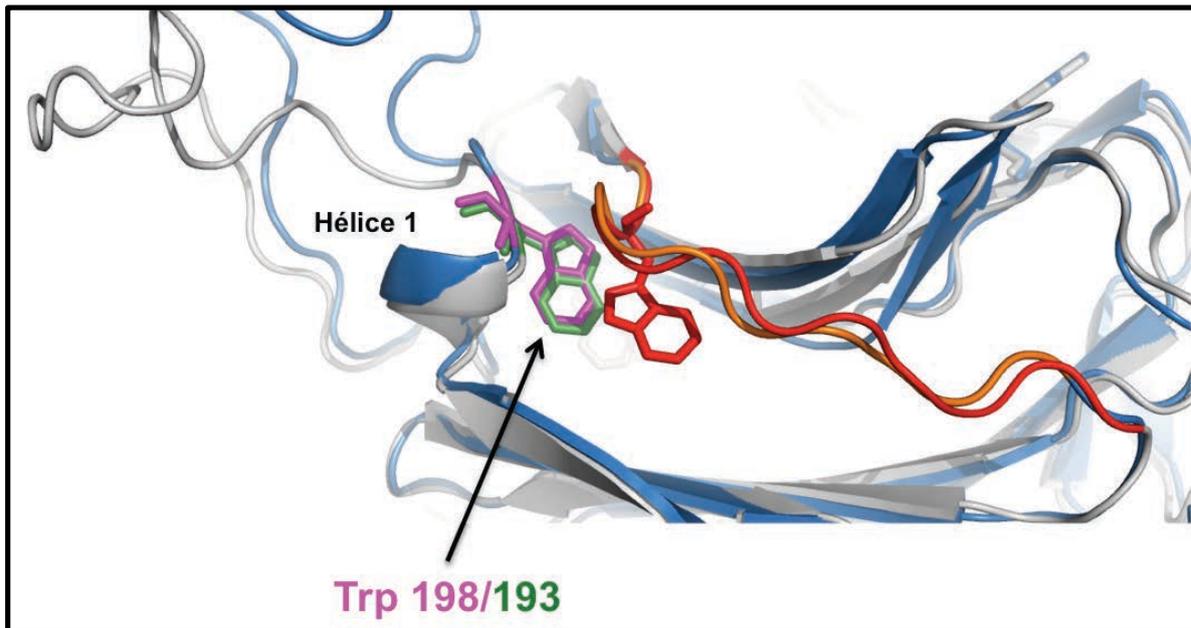


FIGURE 4.III.6 : SITE DE FIXATION DE C1Q DANS LA EHCRT ET LA TCRT

Le site de fixation de C1q suggéré par Paul Eggleton pour la HuCRT, représenté en rouge pour la **TcCRT (EWSKRIRFPEDK)** et en orange pour la **EhCRT (LIRKEIKCESDD)**, se situe à proximité d'un résidu tryptophane conservé (198 pour la TcCRT et 193 pour la EhCRT) à proximité de l'hélice précédant la jonction du domaine P.

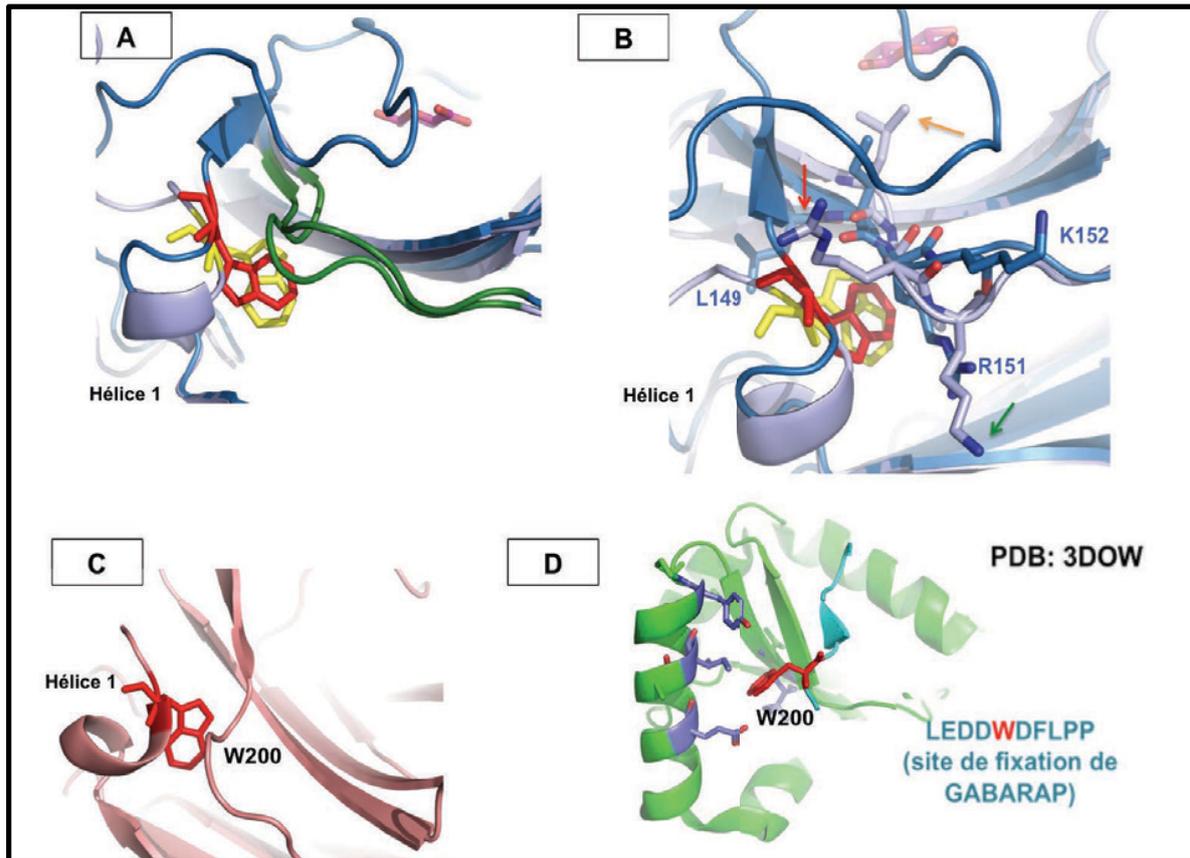


FIGURE 4.III.7 : DEPLACEMENT DU TRYPTOPHANE 198/193 LORS D'UNE INTERACTION

(A) La superposition des chaînes B et C de la EhCRT montre un déplacement du résidu tryptophane de la chaîne C (**rouge**) vers le solvant lors de la fermeture du domaine P. (B) Dans cette même chaîne, les résidus L149, R151 et K152 appartenant à la « séquence Eggleton » basculent de 180° et influencent l'environnement du tryptophane (flèches orange, rouge et verte respectivement). Le glucose à proximité (**magenta**) pourrait ralentir ce basculement. (C) Dans la HuCRT le résidu tryptophane 200 est enfoui vers le domaine globulaire comme dans la EhCRT et la TcCRT. (D) L'utilisation d'un peptide correspondant à la séquence du site de fixation de la protéine GABARAP (**LEDDWDFLPP**) suggère que ce tryptophane 200 participe à l'interaction de la HuCRT avec le ligand et doit donc basculer vers l'extérieur.

L'encadré de la **figure 4.III.8** présente un modèle dans lequel les GRC1q interagiraient avec les calréticulines au niveau de la séquence suggérée par Paul Eggleton, avec la participation des résidus R108-109^B et K200-201^A de C1q. Cette interaction pourrait conduire à une fermeture du domaine P, modifiant la structure de la jonction et conduisant à l'exposition du résidu tryptophane. Cette fermeture du domaine P pourrait correspondre à l'hypothèse de réarrangement conformationnel que nous avons utilisé pour ajuster les courbes d'interactions lors de nos expériences SPR, et conduire à un complexe plus stable. Cette interaction a également été modélisée par Yadav et ses collaborateurs (**annexe 2**) (Yadav et al., 2014).

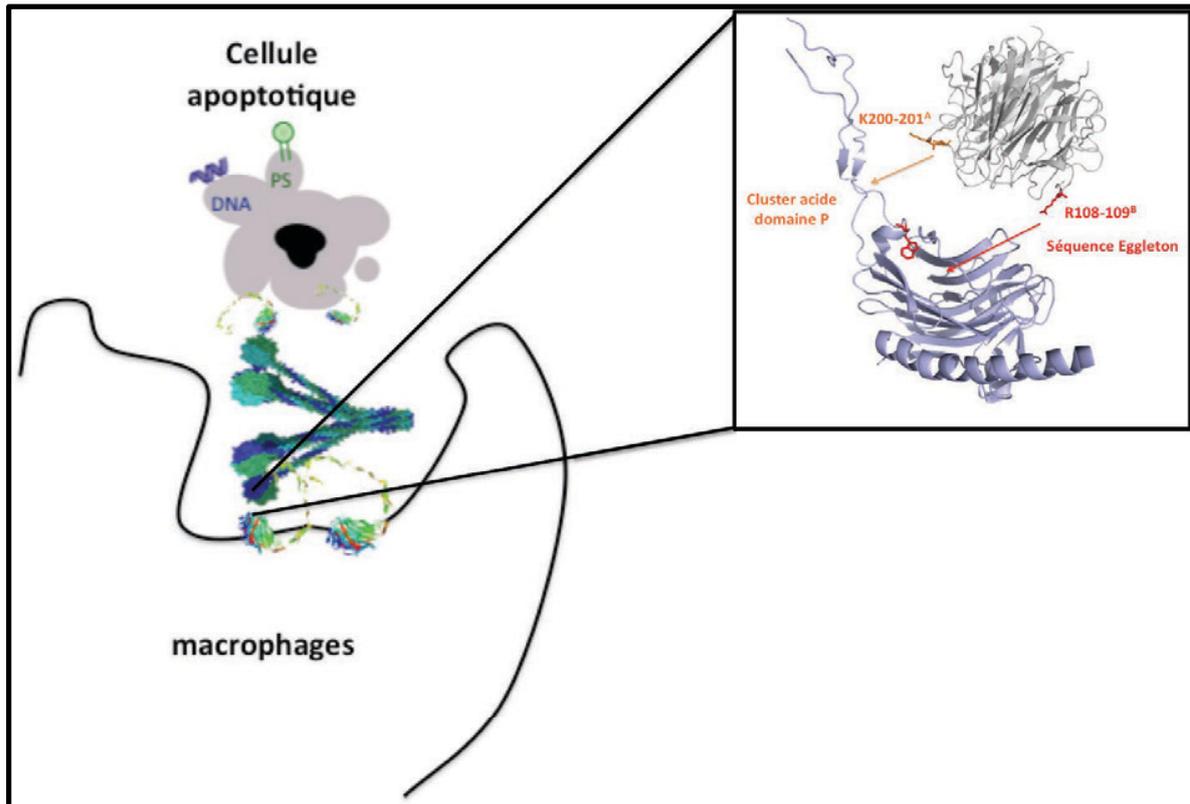


FIGURE 4.III.8 : HYPOTHESE PROPOSEE POUR L'INTERACTION ENTRE C1q ET LA CRT A L'INTERFACE ENTRE LES MACROPHAGES ET LES CELLULES APOPTOTIQUES

Nous pouvons imaginer le scénario suivant : C1q joue le rôle de molécule pontante entre les cellules apoptotiques et les macrophages dans une position allongée et par l'intermédiaire d'une interaction entre la CRT à la surface des deux types cellulaires. Cette interaction implique le domaine globulaire des deux protéines. C1q interagit au niveau d'une séquence LINKDIRCK potentiellement par l'intermédiaire des résidus R108-109^B. L'interaction provoque la fermeture du domaine P, dont la face interne est négative et interagit possiblement avec les résidus K200-201^A de C1q, et l'exposition du W198/193 ou 200 des CRT (encadré).

Afin de confirmer ces hypothèses et pour décrire plus précisément les interactions à un niveau moléculaire entre les CRT et C1q, les études structurales du complexe entre les GRC1q plasmatiques et la TcCRT par RMN seront poursuivies. En cas de succès elles seront appliquées aux HuCRT et EhCRT pour mettre en évidence des sites d'interactions similaires ou différents. En parallèle, la caractérisation du complexe par cristallographie aux rayons X n'est pas abandonnée. Nous pourrions en effet utiliser des peptides correspondant à la séquence potentielle impliquée dans l'interaction avec C1q, de la même manière que cela a été fait pour la protéine GABARAP. Ces études pourraient confirmer le rôle du tryptophane et potentiellement d'autres résidus à proximité. L'implication de ce résidu peut être confirmée par l'utilisation de la spectroscopie de fluorescence. Cette option avait été délaissée, car nous avons montré par la même méthode que tous les tryptophanes étaient enfouis dans la protéine

et de ce fait la présence du partenaire ne permettait pas d'enfouir d'avantage le résidu et modifier son environnement. Si l'interaction expose réellement un tryptophane à la surface, alors il serait alors possible de mesurer un signal différent et d'identifier le résidu par des combinaisons de mutations.

Ces mesures pourront être complétées avec d'autres types d'expériences comme la thermophorèse ou l'ITC (même si l'instabilité de C1q au dessus de 2 μ M limite cette technique) pour mesurer des affinités en solution, plus compatibles avec les études structurales, mais aussi pour confirmer la stoechiométrie de l'interaction et notre modèle 1 :1. Pour confirmer le site de reconnaissance de C1q, les expériences de compétition avec les immunoglobulines déjà réalisées par Yadav et ses collaborateurs (**Yadav et al., 2014**) pourront être répétées sur les Hu, Tc et EhCRT. Dans le cas où l'identification des résidus clés serait une réussite lors des études structurales, nous envisageons de confirmer ces observations en reproduisant les études d'interactions sur différents mutants de C1q et des CRT. L'identification de ces résidus pourrait être couplée à des analyses de type SAXS afin de modéliser l'interaction. L'observation des courbes de dissociation du complexe nous fait néanmoins craindre que la dissociation soit trop rapide pour que le complexe soit stable pour ce type d'études.

Enfin, nous continuerons à travailler sur les GRC1q recombinantes pour obtenir un outil fonctionnel. En plus des vérifications déjà en cours pour comprendre l'absence d'interaction, il est prévu d'effectuer une nouvelle tentative de production en transfectant les cellules HEK-293F successivement par trois plasmides contenant les trois chaînes séparément comme cela a été réalisé avec succès pour C1q (**Bally et al., 2013**).

5. Collaborations et perspectives fonctionnelles

Nous envisageons d'observer l'effet de nos différents fragments et mutants sur l'infectivité et les mécanismes d'évasion du système immunitaire de *T. cruzi* en collaboration avec le groupe du Dr. Ferreira. En effet, si le domaine globulaire des CRTs est suffisant pour l'interaction avec C1q, il se pourrait que d'autres domaines puissent intervenir dans les mécanismes d'infection et d'inhibition. Lors de ma thèse, j'ai pu débiter des expériences dans le laboratoire du Dr. Ferreira pendant un séjour d'un mois au Chili. Pendant cette période, nous avons souhaité vérifier l'effet des troncations du domaine P sur l'inhibition de la voie classique du complément et l'infectivité du parasite. De premières expériences, qui devront être répétées, ont suggéré que même en absence de domaine P, la TcCRT continuait à inhiber

la voie classique du complément. Contrairement au rôle du domaine P décrit par le Dr. Ferreira (**Valck et al., 2010**) dans l'inhibition par interaction avec les protéases C1r et C1s, il se pourrait que la TcCRT inhibe le complément en masquant simplement la surface de reconnaissance des IgG ou des IgM (le mutant arginine de C1q inhibant l'interaction avec les IgM, IgG et l'activation de la voie classique du complément). Nous avons voulu reproduire ces expériences sur les autres calréticulines, mais il a été extrêmement difficile d'importer le protocole utilisé à Santiago dans notre laboratoire. Après de nombreuses mises au point, j'ai pu établir un protocole permettant d'avoir des contrôles positifs et négatifs fiables présentés dans la section Matériels et Méthodes, mais je n'ai jamais pu observer d'effet inhibiteur des trois CRTs sur l'activation de la voie classique du complément.

Le Dr. Ferreira a également mis au point un protocole pour mettre en évidence le rôle de la CRT dans l'infectivité. Pour cela, la TcCRT recombinante est fixée à la surface de la forme non infectieuse épimastigote de *T. cruzi* pour la rendre infectieuse. Lors de ma visite à Santiago, nous avons essayé de définir un protocole permettant de tester la fixation de mes différents fragments à la surface de l'épimastigote, pour ensuite vérifier leurs effets sur la restauration de l'infectivité. Ces expériences très préliminaires seront poursuivies à Santiago pour identifier le domaine de la TcCRT responsable de l'infectivité et potentiellement l'effet de futurs mutants.

La stratégie déployée pour déchiffrer les interactions entre les calréticulines et C1q, et l'effet de ces protéines sur l'action pathogène des parasites, pourra être extrapolée à d'autres micro-organismes utilisant le même type de stratégie mimétique. C'est notamment le cas de la bactérie *Streptococcus Pneumoniae*, exprimant à sa surface une GAPDH reconnue par C1q et conduisant à l'activation de la voie classique du complément. Comme pour les parasites que nous avons étudiés, la GAPDH est présente à la surface des cellules apoptotiques (**Terrasse et al., 2012**). Dans ce contexte, j'ai pu résoudre la structure de la GAPDH du pneumocoque sous forme de dimère et de tétramère à des résolutions maximales de 3,3 Å et 1,9 Å, respectivement, en collaboration avec Anne-Marie Di Guilmi du groupe PG à l'IBS. Cette structure a permis de vérifier et identifier des sites potentiels d'interaction à la surface de la GAPDH pour C1q et le plasminogène (autre molécule de l'hôte détournée par le pneumocoque grâce à la GAPDH et nécessaire à l'infection) et pourra servir de base pour poursuivre la compréhension des mécanismes de subversion de cette bactérie.

BIBLIOGRAPHIE

A

- Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.-W., Kapral, G.J., Grosse-Kunstleve, R.W., et al. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* *66*, 213–221.
- Agrawal, A., Shrive, A.K., Greenhough, T.J., and Volanakis, J.E. (2001). Topology and structure of the C1q-binding site on C-reactive protein. *J. Immunol. Baltim. Md 1950* *166*, 3998–4004.
- Aguilar, L., Ramírez, G., Valck, C., Molina, M.C., Rojas, A., Schwaeble, W., Ferreira, V., and Ferreira, A. (2005). F(ab')₂ antibody fragments against *Trypanosoma cruzi* calreticulin inhibit its interaction with the first component of human complement. *Biol. Res.* *38*, 187–195.
- Aguilar-Guzmán, L., Lobos-González, L., Rosas, C., Vallejos, G., Falcón, C., Sosoniuk, E., Coddou, F., Leyton, L., Lemus, D., Quest, A.F.G., et al. (2014). Human survivin and *Trypanosoma cruzi* calreticulin act in synergy against a murine melanoma in vivo. *PloS One* *9*, e95457.
- Aguillón, J.C., Ferreira, L., Pérez, C., Colombo, A., Molina, M.C., Wallace, A., Solari, A., Carvallo, P., Galindo, M., Galanti, N., et al. (2000). Tc45, a dimorphic *Trypanosoma cruzi* immunogen with variable chromosomal localization, is calreticulin. *Am. J. Trop. Med. Hyg.* *63*, 306–312.
- Aksamit, R.R., Falk, W., and Leonard, E.J. (1981). Chemotaxis by mouse macrophage cell lines. *J. Immunol. Baltim. Md 1950* *126*, 2194–2199.
- Albertí, S., Marqués, G., Hernández-Allés, S., Rubires, X., Tomás, J.M., Vivanco, F., and Benedí, V.J. (1996). Interaction between complement subcomponent C1q and the *Klebsiella pneumoniae* porin OmpK36. *Infect. Immun.* *64*, 4719–4725.
- Alder, N.N., Shen, Y., Brodsky, J.L., Hendershot, L.M., and Johnson, A.E. (2005). The molecular mechanisms underlying BiP-mediated gating of the Sec61 translocon of the endoplasmic reticulum. *J. Cell Biol.* *168*, 389–399.
- Aleshin, A.E., Schraufstatter, I.U., Stec, B., Bankston, L.A., Liddington, R.C., and DiScipio, R.G. (2012a). Structure of complement C6 suggests a mechanism for initiation and unidirectional, sequential assembly of membrane attack complex (MAC). *J. Biol. Chem.* *287*, 10210–10222.
- Aleshin, A.E., DiScipio, R.G., Stec, B., and Liddington, R.C. (2012b). Crystal structure of C5b-6 suggests structural basis for priming assembly of the membrane attack complex. *J. Biol. Chem.* *287*, 19642–19652.
- Abd-Alla, M.D., Jackson, T.F.G.H., Rogers, T., Reddy, S., and Ravdin, J.I. (2006). Mucosal immunity to asymptomatic *Entamoeba histolytica* and *Entamoeba dispar* infection is associated with a peak intestinal anti-lectin immunoglobulin A antibody response. *Infect. Immun.* *74*, 3897–3903.
- Amara, U., Rittirsch, D., Flierl, M., Bruckner, U., Klos, A., Gebhard, F., Lambris, J.D., and Huber-Lang, M. (2008). Interaction between the coagulation and complement system. *Adv. Exp. Med. Biol.* *632*, 71–79.

- Amara, U., Flierl, M.A., Rittirsch, D., Klos, A., Chen, H., Acker, B., Brückner, U.B., Nilsson, B., Gebhard, F., Lambris, J.D., et al. (2010). Molecular intercommunication between the complement and coagulation systems. *J. Immunol. Baltim. Md 1950* *185*, 5628–5636.
- Andrin, C., Corbett, E.F., Johnson, S., Dabrowska, M., Campbell, I.D., Eggleton, P., Opas, M., and Michalak, M. (2000). Expression and purification of mammalian calreticulin in *Pichia pastoris*. *Protein Expr. Purif.* *20*, 207–215.
- Anfinsen, C.B. (1973). Principles that govern the folding of protein chains. *Science* *181*, 223–230.
- Anfinsen, C.B., and Haber, E. (1961). Studies on the reduction and re-formation of protein disulfide bonds. *J. Biol. Chem.* *236*, 1361–1363.
- Anfinsen, C.B., Haber, E., Sela, M., and White, F.H. (1961). The kinetics of formation of native ribonuclease during oxidation of the reduced polypeptide chain. *Proc. Natl. Acad. Sci. U. S. A.* *47*, 1309–1314.
- Araya, J.E., Cornejo, A., Orrego, P.R., Cordero, E.M., Cortéz, M., Olivares, H., Neira, I., Sagua, H., da Silveira, J.F., Yoshida, N., et al. (2008). Calcineurin B of the human protozoan parasite *Trypanosoma cruzi* is involved in cell invasion. *Microbes Infect. Inst. Pasteur* *10*, 892–900.
- Arlaud, G.J., Sim, R.B., Duplaa, A.M., and Colomb, M.G. (1979). Differential elution of Clq,Clr and Cls from human C1 bound to immune aggregates. Use in the rapid purification of C1 subcomponents. *Mol. Immunol.* *16*, 445–450.
- Arlaud, G.J., Colomb, M.G., and Gagnon, J. (1987). A functional model of the human C1 complex Emergence of a functional model. *Immunol. Today* *8*, 106–111.
- Arlaud, G.J., Gaboriaud, C., Thielens, N.M., Rossi, V., Bersch, B., Hernandez, J.F., and Fontecilla-Camps, J.C. (2001). Structural biology of C1: dissection of a complex molecular machinery. *Immunol. Rev.* *180*, 136–145.
- Arlaud, G.J., Gaboriaud, C., Garnier, G., Circolo, A., Thielens, N.M., Budayova-Spano, M., Fontecilla-Camps, J.C., and Volanakis, J.E. (2002a). Structure, function and molecular genetics of human and murine C1r. *Immunobiology* *205*, 365–382.
- Arlaud, G.J., Gaboriaud, C., Thielens, N.M., Budayova-Spano, M., Rossi, V., and Fontecilla-Camps, J.C. (2002b). Structural biology of the C1 complex of complement unveils the mechanisms of its activation and proteolytic activity. *Mol. Immunol.* *39*, 383–394.
- Arnaudeau, S., Frieden, M., Nakamura, K., Castelbou, C., Michalak, M., and Demaurex, N. (2002). Calreticulin differentially modulates calcium uptake and release in the endoplasmic reticulum and mitochondria. *J. Biol. Chem.* *277*, 46696–46705.
- Aslund, L., Carlsson, L., Henriksson, J., Rydåker, M., Toro, G.C., Galanti, N., and Pettersson, U. (1994). A gene family encoding heterogeneous histone H1 proteins in *Trypanosoma cruzi*. *Mol. Biochem. Parasitol.* *65*, 317–330.
- Ayo, C.M., Dalalio, M.M. de O., Visentainer, J.E.L., Reis, P.G., Sippert, E. ., Jarduli, L.R., Alves, H.V., and Sell, A.M. (2013). Genetic susceptibility to Chagas disease: an overview about the infection and about the association between disease and the immune response genes. *BioMed Res. Int.* *2013*, 284729.

B

- Baksh, S., and Michalak, M. (1991). Expression of calreticulin in *Escherichia coli* and identification of its Ca²⁺ binding domains. *J. Biol. Chem.* *266*, 21458–21465.
- Baksh, S., Spamer, C., Heilmann, C., and Michalak, M. (1995). Identification of the Zn²⁺ binding region in calreticulin. *FEBS Lett.* *376*, 53–57.
- Bally, I., Rossi, V., Lunardi, T., Thielens, N.M., Gaboriaud, C., and Arlaud, G.J. (2009). Identification of the C1q-binding Sites of Human C1r and C1s: a refined three-dimensional model of the C1 complex of complement. *J. Biol. Chem.* *284*, 19340–19348.
- Bally, I., Ancelet, S., Moriscot, C., Gonnet, F., Mantovani, A., Daniel, R., Schoehn, G., Arlaud, G.J., and Thielens, N.M. (2013). Expression of recombinant human complement C1q allows identification of the C1r/C1s-binding sites. *Proc. Natl. Acad. Sci. U. S. A.* *110*, 8650–8655.
- Band, R.N., and Cirrito, H. (1979). Growth response of axenic *Entamoeba histolytica* to hydrogen, carbon dioxide, and oxygen. *J. Protozool.* *26*, 282–286.
- Banerjee, S., Vishwanath, P., Cui, J., Kelleher, D.J., Gilmore, R., Robbins, P.W., and Samuelson, J. (2007). The evolution of N-glycan-dependent endoplasmic reticulum quality control factors for glycoprotein folding and degradation. *Proc. Natl. Acad. Sci. U. S. A.* *104*, 11676–11681.
- Bansal, D., Ave, P., Kerneis, S., Frileux, P., Boché, O., Baglin, A.C., Dubost, G., Leguern, A.-S., Prevost, M.-C., Bracha, R., et al. (2009). An ex-vivo human intestinal model to study *Entamoeba histolytica* pathogenesis. *PLoS Negl. Trop. Dis.* *3*, e551.
- Basso, B. (2013). Modulation of immune response in experimental Chagas disease. *World J. Exp. Med.* *3*, 1–10.
- Bastianutto, C., Clementi, E., Codazzi, F., Podini, P., De Giorgi, F., Rizzuto, R., Meldolesi, J., and Pozzan, T. (1995). Overexpression of calreticulin increases the Ca²⁺ capacity of rapidly exchanging Ca²⁺ stores and reveals aspects of their luminal microenvironment and function. *J. Cell Biol.* *130*, 847–855.
- Basu, S., Binder, R.J., Ramalingam, T., and Srivastava, P.K. (2001). CD91 is a common receptor for heat shock proteins gp96, hsp90, hsp70, and calreticulin. *Immunity* *14*, 303–313.
- Berasain, P., Carmona, C., Frangione, B., Cazzulo, J.J., and Goñi, F. (2003). Specific cleavage sites on human IgG subclasses by cruzipain, the major cysteine proteinase from *Trypanosoma cruzi*. *Mol. Biochem. Parasitol.* *130*, 23–29.
- Bhakdi, S., and Trantum-Jensen, J. (1991). Complement lysis: a hole is a hole. *Immunol. Today* *12*, 318–320; discussion 321.
- Bobak, D.A., Gaither, T.A., Frank, M.M., and Tenner, A.J. (1987). Modulation of FcR function by complement: subcomponent C1q enhances the phagocytosis of IgG-opsonized targets by human monocytes and culture-derived macrophages. *J. Immunol.* *Baltim. Md 1950* *138*, 1150–1156.

- Bobak, D.A., Frank, M.M., and Tenner, A.J. (1988). C1q acts synergistically with phorbol dibutyrate to activate CR1-mediated phagocytosis by human mononuclear phagocytes. *Eur. J. Immunol.* *18*, 2001–2007.
- Bohlon, S.S., Fraser, D.A., and Tenner, A.J. (2007). Complement proteins C1q and MBL are pattern recognition molecules that signal immediate and long-term protective immune functions. *Mol. Immunol.* *44*, 33–43.
- Bontempi, E.J., Búa, J., Aslund, L., Porcel, B., Segura, E.L., Henriksson, J., Orn, A., Pettersson, U., and Ruiz, A.M. (1993). Isolation and characterization of a gene from *Trypanosoma cruzi* encoding a 46-kilodalton protein with homology to human and rat tyrosine aminotransferase. *Mol. Biochem. Parasitol.* *59*, 253–262.
- Botto, M., Dell’Agnola, C., Bygrave, A.E., Thompson, E.M., Cook, H.T., Petry, F., Loos, M., Pandolfi, P.P., and Walport, M.J. (1998). Homozygous C1q deficiency causes glomerulonephritis associated with multiple apoptotic bodies. *Nat. Genet.* *19*, 56–59.
- Braakman, I., and Balleid, N.J. (2011). Protein folding and modification in the mammalian endoplasmic reticulum. *Annu. Rev. Biochem.* *80*, 71–99.
- Braga, L.L., Ninomiya, H., McCoy, J.J., Eacker, S., Wiedmer, T., Pham, C., Wood, S., Sims, P.J., and Petri, W.A. (1992). Inhibition of the complement membrane attack complex by the galactose-specific adhesion of *Entamoeba histolytica*. *J. Clin. Invest.* *90*, 1131–1137.
- Bramham, J., Thai, C.-T., Soares, D.C., Uhrin, D., Ogata, R.T., and Barlow, P.N. (2005). Functional insights from the structure of the multifunctional C345C domain of C5 of complement. *J. Biol. Chem.* *280*, 10636–10645.
- Brockmeier, A., and Williams, D.B. (2006). Potent lectin-independent chaperone function of calnexin under conditions prevalent within the lumen of the endoplasmic reticulum. *Biochemistry (Mosc.)* *45*, 12906–12916.
- Bruchhaus, I., and Tannich, E. (1994). Induction of the iron-containing superoxide dismutase in *Entamoeba histolytica* by a superoxide anion-generating system or by iron chelation. *Mol. Biochem. Parasitol.* *67*, 281–288.
- Brünger, A.T. (1992). Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* *355*, 472–475.
- Budayova-Spano, M., Lacroix, M., Thielens, N.M., Arlaud, G.J., Fontecilla-Camps, J.C., and Gaboriaud, C. (2002a). The crystal structure of the zymogen catalytic domain of complement protease C1r reveals that a disruptive mechanical stress is required to trigger activation of the C1 complex. *EMBO J.* *21*, 231–239.
- Budayova-Spano, M., Grabarse, W., Thielens, N.M., Hillen, H., Lacroix, M., Schmidt, M., Fontecilla-Camps, J.C., Arlaud, G.J., and Gaboriaud, C. (2002b). Monomeric structures of the zymogen and active catalytic domain of complement protease c1r: further insights into the c1 activation mechanism. *Struct. Lond. Engl.* *1993* *10*, 1509–1519.
- Burton, D.R., Boyd, J., Brampton, A.D., Easterbrook-Smith, S.B., Emanuel, E.J., Novotny, J., Rademacher, T.W., van Schravendijk, M.R., Sternberg, M.J., and Dwek, R.A. (1980). The C1q receptor site on immunoglobulin G. *Nature* *288*, 338–344.

C

- Cabral, C.M., Liu, Y., Moremen, K.W., and Sifers, R.N. (2002). Organizational diversity among distinct glycoprotein endoplasmic reticulum-associated degradation programs. *Mol. Biol. Cell* *13*, 2639–2650.
- Cai, K.X., Tse, L.Y., Leung, C., Tam, P.K.H., Xu, R., and Sham, M.H. (2008). Suppression of lung tumor growth and metastasis in mice by adeno-associated virus-mediated expression of vasostatin. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* *14*, 939–949.
- Campetella, O., Sánchez, D., Cazzulo, J.J., and Frasch, A.C. (1992). A superfamily of *Trypanosoma cruzi* surface antigens. *Parasitol. Today Pers. Ed* *8*, 378–381.
- Caramelo, J.J., Castro, O.A., de Prat-Gay, G., and Parodi, A.J. (2004). The endoplasmic reticulum glucosyltransferase recognizes nearly native glycoprotein folding intermediates. *J. Biol. Chem.* *279*, 46280–46285.
- Cardenas, T.C., Johnson, C.A., Pratap, S., Nde, P.N., Furtak, V., Kleshchenko, Y.Y., Lima, M.F., and Villalta, F. (2010). REGULATION of the EXTRACELLULAR MATRIX INTERACTOME by *Trypanosoma cruzi*. *Open Parasitol. J.* *4*, 72–76.
- Carroll, M.V., and Sim, R.B. (2011). Complement in health and disease. *Adv. Drug Deliv. Rev.* *63*, 965–975.
- Castillo, C., Ramírez, G., Valck, C., Aguilar, L., Maldonado, I., Rosas, C., Galanti, N., Kemmerling, U., and Ferreira, A. (2013). The interaction of classical complement component C1 with parasite and host calreticulin mediates *Trypanosoma cruzi* infection of human placenta. *PLoS Negl. Trop. Dis.* *7*, e2376.
- Chadee, K., Petri, W.A., Innes, D.J., and Ravdin, J.I. (1987). Rat and human colonic mucins bind to and inhibit adherence lectin of *Entamoeba histolytica*. *J. Clin. Invest.* *80*, 1245–1254.
- Chao, M.P., Jaiswal, S., Weissman-Tsukamoto, R., Alizadeh, A.A., Gentles, A.J., Volkmer, J., Weiskopf, K., Willingham, S.B., Raveh, T., Park, C.Y., et al. (2010). Calreticulin is the dominant pro-phagocytic signal on multiple human cancers and is counterbalanced by CD47. *Sci. Transl. Med.* *2*, 63ra94.
- Chávez-Munguía, B., Talamás-Rohana, P., Castañón, G., Salazar-Villatoro, L., Hernández-Ramírez, V., and Martínez-Palomo, A. (2012). Differences in cap formation between invasive *Entamoeba histolytica* and non-invasive *Entamoeba dispar*. *Parasitol. Res.* *111*, 215–221.
- Choi, M.-H., Sajed, D., Poole, L., Hirata, K., Herdman, S., Torian, B.E., and Reed, S.L. (2005). An unusual surface peroxiredoxin protects invasive *Entamoeba histolytica* from oxidant attack. *Mol. Biochem. Parasitol.* *143*, 80–89.
- Chouquet, A., Païdassi, H., Ling, W.L., Frachet, P., Houen, G., Arlaud, G.J., and Gaboriaud, C. (2011). X-ray structure of the human calreticulin globular domain reveals a peptide-binding area and suggests a multi-molecular mechanism. *PloS One* *6*, e17886.

Del Cid, N., Jeffery, E., Rizvi, S.M., Stamper, E., Peters, L.R., Brown, W.C., Provoda, C., and Raghavan, M. (2010). Modes of calreticulin recruitment to the major histocompatibility complex class I assembly pathway. *J. Biol. Chem.* *285*, 4520–4535.

Comis, A., and Easterbrook-Smith, S.B. (1985). Evidence for arginine residues in the immunoglobulin-binding sites of human Clq. *Biochim. Biophys. Acta* *842*, 45–51.

Conte, I., Labriola, C., Cazzulo, J.J., Docampo, R., and Parodi, A.J. (2003). The interplay between folding-facilitating mechanisms in *Trypanosoma cruzi* endoplasmic reticulum. *Mol. Biol. Cell* *14*, 3529–3540.

Corbett, E.F., Michalak, K.M., Oikawa, K., Johnson, S., Campbell, I.D., Eggleton, P., Kay, C., and Michalak, M. (2000). The conformation of calreticulin is influenced by the endoplasmic reticulum luminal environment. *J. Biol. Chem.* *275*, 27177–27185.

Coura, J.R., and Borges-Pereira, J. (2012). Chagas disease. What is known and what should be improved: a systemic review. *Rev. Soc. Bras. Med. Trop.* *45*, 286–296.

D’Cruz, D.P., Khamashta, M.A., and Hughes, G.R.V. (2007). Systemic lupus erythematosus. *Lancet* *369*, 587–596.

D

Desjardins, M. (2003). ER-mediated phagocytosis: a new membrane for new functions. *Nat. Rev. Immunol.* *3*, 280–291.

Dharampaul, S., Kaufman, W.R., and Belosevic, M. (1993). Differential recognition of saliva antigens from the ixodid tick *Amblyomma hebraeum* (Acari: Ixodidae) by sera from infested and immunized rabbits. *J. Med. Entomol.* *30*, 262–266.

Diebold, C.A., Beurskens, F.J., de Jong, R.N., Koning, R.I., Strumane, K., Lindorfer, M.A., Voorhorst, M., Ugurlar, D., Rosati, S., Heck, A.J.R., et al. (2014). Complement is activated by IgG hexamers assembled at the cell surface. *Science* *343*, 1260–1263.

Diederichs, K., and Karplus, P.A. (2013). Better models by discarding data? *Acta Crystallogr. D Biol. Crystallogr.* *69*, 1215–1222.

Dill, K.A. (1985). Theory for the folding and stability of globular proteins. *Biochemistry (Mosc.)* *24*, 1501–1509.

Donnelly, S., Roake, W., Brown, S., Young, P., Naik, H., Wordsworth, P., Isenberg, D.A., Reid, K.B.M., and Eggleton, P. (2006). Impaired recognition of apoptotic neutrophils by the C1q/calreticulin and CD91 pathway in systemic lupus erythematosus. *Arthritis Rheum.* *54*, 1543–1556.

Duggal, P., Guo, X., Haque, R., Peterson, K.M., Ricklefs, S., Mondal, D., Alam, F., Noor, Z., Verkerke, H.P., Marie, C., et al. (2011). A mutation in the leptin receptor is associated with *Entamoeba histolytica* infection in children. *J. Clin. Invest.* *121*, 1191–1198.

Duus, K., Sandhu, N., Jørgensen, C.S., Hansen, P.R., Steinø, A., Thaysen-Andersen, M., Højrup, P., and Houen, G. (2009). Interaction of the chaperone calreticulin with proteins and peptides of different structural classes. *Protein Pept. Lett.* *16*, 1414–1423.

E

- Easterbrook-Smith, S.B. (1983). Activation of the binding of C1q to immune complexes by zinc. *FEBS Lett.* 162, 117–119.
- Edelhoch, H. (1967). Spectroscopic determination of tryptophan and tyrosine in proteins. *Biochemistry (Mosc.)* 6, 1948–1954.
- Edman, P. (1949). A method for the determination of amino acid sequence in peptides. *Arch. Biochem.* 22, 475.
- Eggleton, P., Lieu, T.S., Zappi, E.G., Sastry, K., Coburn, J., Zaner, K.S., Sontheimer, R.D., Capra, J.D., Ghebrehiwet, B., and Tauber, A.I. (1994). Calreticulin is released from activated neutrophils and binds to C1q and mannan-binding protein. *Clin. Immunol. Immunopathol.* 72, 405–409.
- Ehrengruber, M.U., Geiser, T., and Deranleau, D.A. (1994). Activation of human neutrophils by C3a and C5A. Comparison of the effects on shape changes, chemotaxis, secretion, and respiratory burst. *FEBS Lett.* 346, 181–184.
- El Homsí, M., Ducroc, R., Claustre, J., Jourdan, G., Gertler, A., Estienne, M., Bado, A., Scoazec, J.-Y., and Plaisancié, P. (2007). Leptin modulates the expression of secreted and membrane-associated mucins in colonic epithelial cells by targeting PKC, PI3K, and MAPK pathways. *Am. J. Physiol. Gastrointest. Liver Physiol.* 293, G365–G373.
- Ellgaard, L., Riek, R., Braun, D., Herrmann, T., Helenius, A., and Wüthrich, K. (2001a). Three-dimensional structure topology of the calreticulin P-domain based on NMR assignment. *FEBS Lett.* 488, 69–73.
- Ellgaard, L., Riek, R., Herrmann, T., Güntert, P., Braun, D., Helenius, A., and Wüthrich, K. (2001b). NMR structure of the calreticulin P-domain. *Proc. Natl. Acad. Sci. U. S. A.* 98, 3133–3138.
- Ellgaard, L., Bettendorff, P., Braun, D., Herrmann, T., Fiorito, F., Jelesarov, I., Güntert, P., Helenius, A., and Wüthrich, K. (2002). NMR structures of 36 and 73-residue fragments of the calreticulin P-domain. *J. Mol. Biol.* 322, 773–784.
- Elliott, T., and Williams, A. (2005). The optimization of peptide cargo bound to MHC class I molecules by the peptide-loading complex. *Immunol. Rev.* 207, 89–99.
- Elnekave, K., Siman-Tov, R., and Ankri, S. (2003). Consumption of L-arginine mediated by *Entamoeba histolytica* L-arginase (EhArg) inhibits amoebicidal activity and nitric oxide production by activated macrophages. *Parasite Immunol.* 25, 597–608.
- Elsner, J., Oppermann, M., Czech, W., Dobos, G., Schöpf, E., Norgauer, J., and Kapp, A. (1994a). C3a activates reactive oxygen radical species production and intracellular calcium transients in human eosinophils. *Eur. J. Immunol.* 24, 518–522.
- Elsner, J., Oppermann, M., Czech, W., and Kapp, A. (1994b). C3a activates the respiratory burst in human polymorphonuclear neutrophilic leukocytes via pertussis toxin-sensitive G-proteins. *Blood* 83, 3324–3331.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* *60*, 2126–2132.

Espinosa-Cantellano, M., and Martínez-Palomo, A. (2000). Pathogenesis of intestinal amebiasis: from molecules to disease. *Clin. Microbiol. Rev.* *13*, 318–331.

F

Faggioni, R., Feingold, K.R., and Grunfeld, C. (2001). Leptin regulation of the immune response and the immunodeficiency of malnutrition. *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.* *15*, 2565–2571.

Favier, A., and Brutscher, B. (2011). Recovering lost magnetization: polarization enhancement in biomolecular NMR. *J. Biomol. NMR* *49*, 9–15.

Fearon, D.T., and Austen, K.F. (1975). Properdin: binding to C3b and stabilization of the C3b-dependent C3 convertase. *J. Exp. Med.* *142*, 856–863.

Ferreira, V., Valck, C., Sánchez, G., Gingras, A., Tzima, S., Molina, M.C., Sim, R., Schwaeble, W., and Ferreira, A. (2004a). The classical activation pathway of the human complement system is specifically inhibited by calreticulin from *Trypanosoma cruzi*. *J. Immunol. Baltim. Md 1950* *172*, 3042–3050.

Ferreira, V., Molina, M.C., Valck, C., Rojas, A., Aguilar, L., Ramírez, G., Schwaeble, W., and Ferreira, A. (2004b). Role of calreticulin from parasites in its interaction with vertebrate hosts. *Mol. Immunol.* *40*, 1279–1291.

Ferreira, V., Molina, M.C., Schwaeble, W., Lemus, D., and Ferreira, A. (2005). Does *Trypanosoma cruzi* calreticulin modulate the complement system and angiogenesis? *Trends Parasitol.* *21*, 169–174.

Fliegel, L., Burns, K., MacLennan, D.H., Reithmeier, R.A., and Michalak, M. (1989). Molecular cloning of the high affinity calcium-binding protein (calreticulin) of skeletal muscle sarcoplasmic reticulum. *J. Biol. Chem.* *264*, 21522–21528.

Fraser, D.A., Laust, A.K., Nelson, E.L., and Tenner, A.J. (2009). C1q differentially modulates phagocytosis and cytokine responses during ingestion of apoptotic cells by human monocytes, macrophages, and dendritic cells. *J. Immunol. Baltim. Md 1950* *183*, 6175–6185.

Fraser, D.A., Pisalyaput, K., and Tenner, A.J. (2010). C1q enhances microglial clearance of apoptotic neurons and neuronal blebs, and modulates subsequent inflammatory cytokine production. *J. Neurochem.* *112*, 733–743.

Frickel, E.-M., Riek, R., Jelesarov, I., Helenius, A., Wuthrich, K., and Ellgaard, L. (2002). TROSY-NMR reveals interaction between ERp57 and the tip of the calreticulin P-domain. *Proc. Natl. Acad. Sci. U. S. A.* *99*, 1954–1959.

G

- Gaboriaud, C., Rossi, V., Bally, I., Arlaud, G.J., and Fontecilla-Camps, J.C. (2000). Crystal structure of the catalytic domain of human complement c1s: a serine protease with a handle. *EMBO J.* *19*, 1755–1765.
- Gaboriaud, C., Juanhuix, J., Gruez, A., Lacroix, M., Darnault, C., Pignol, D., Verger, D., Fontecilla-Camps, J.C., and Arlaud, G.J. (2003). The crystal structure of the globular head of complement protein C1q provides a basis for its versatile recognition properties. *J. Biol. Chem.* *278*, 46974–46982.
- Gaboriaud, C., Thielens, N.M., Gregory, L.A., Rossi, V., Fontecilla-Camps, J.C., and Arlaud, G.J. (2004). Structure and activation of the C1 complex of complement: unraveling the puzzle. *Trends Immunol.* *25*, 368–373.
- Gaboriaud, C., Frachet, P., Thielens, N.M., and Arlaud, G.J. (2011a). The human c1q globular domain: structure and recognition of non-immune self ligands. *Front. Immunol.* *2*, 92.
- Gaboriaud, C., Gregory-Pauron, L., Teillet, F., Thielens, N.M., Bally, I., and Arlaud, G.J. (2011b). Structure and properties of the Ca(2+)-binding CUB domain, a widespread ligand-recognition unit involved in major biological functions. *Biochem. J.* *439*, 185–193.
- Gaboriaud, C., Ling, W.L., Thielens, N.M., Bally, I., and Rossi, V. (2014). Deciphering the fine details of c1 assembly and activation mechanisms: “mission impossible”? *Front. Immunol.* *5*, 565.
- Gagnon, E., Duclos, S., Rondeau, C., Chevet, E., Cameron, P.H., Steele-Mortimer, O., Paiement, J., Bergeron, J.J.M., and Desjardins, M. (2002). Endoplasmic reticulum-mediated phagocytosis is a mechanism of entry into macrophages. *Cell* *110*, 119–131.
- Galvan, M.D., Greenlee-Wacker, M.C., and Bohlsion, S.S. (2012). C1q and phagocytosis: the perfect complement to a good meal. *J. Leukoc. Biol.* *92*, 489–497.
- Galván-Moroyoqui, J.M., Del Carmen Domínguez-Robles, M., and Meza, I. (2011). Pathogenic bacteria prime the induction of Toll-like receptor signalling in human colonic cells by the Gal/GalNAc lectin Carbohydrate Recognition Domain of *Entamoeba histolytica*. *Int. J. Parasitol.* *41*, 1101–1112.
- Ganoza, M.C., and Williams, C.A. (1969). In vitro synthesis of different categories of specific protein by membrane-bound and free ribosomes. *Proc. Natl. Acad. Sci. U. S. A.* *63*, 1370–1376.
- Gardai, S.J., Xiao, Y.-Q., Dickinson, M., Nick, J.A., Voelker, D.R., Greene, K.E., and Henson, P.M. (2003). By binding SIRPalpha or calreticulin/CD91, lung collectins act as dual function surveillance molecules to suppress or enhance inflammation. *Cell* *115*, 13–23.
- Gardai, S.J., McPhillips, K.A., Frasch, S.C., Janssen, W.J., Starefeldt, A., Murphy-Ullrich, J.E., Bratton, D.L., Oldenborg, P.-A., Michalak, M., and Henson, P.M. (2005). Cell-surface calreticulin initiates clearance of viable or apoptotic cells through trans-activation of LRP on the phagocyte. *Cell* *123*, 321–334.
- Garlatti, V., Chouquet, A., Lunardi, T., Vivès, R., Païdassi, H., Lortat-Jacob, H., Thielens, N.M., Arlaud, G.J., and Gaboriaud, C. (2010). Cutting edge: C1q binds deoxyribose and heparan sulfate through neighboring sites of its recognition domain. *J. Immunol. Baltim. Md 1950* *185*, 808–812.

- Ge, H., Xiong, Y., Lemon, B., Lee, K.J., Tang, J., Wang, P., Weiszmann, J., Hawkins, N., Laudemann, J., Min, X., et al. (2010). Generation of novel long-acting globular adiponectin molecules. *J. Mol. Biol.* *399*, 113–119.
- Geiser, M., Cèbe, R., Drewello, D., and Schmitz, R. (2001). Integration of PCR fragments at any specific site within cloning vectors without the use of restriction enzymes and DNA ligase. *BioTechniques* *31*, 88–90, 92.
- Ghai, R., Waters, P., Roumenina, L.T., Gadjeva, M., Kojouharova, M.S., Reid, K.B.M., Sim, R.B., and Kishore, U. (2007). C1q and its growing family. *Immunobiology* *212*, 253–266.
- Gilmore, R., Walter, P., and Blobel, G. (1982). Protein translocation across the endoplasmic reticulum. II. Isolation and characterization of the signal recognition particle receptor. *J. Cell Biol.* *95*, 470–477.
- Goicoechea, S., Pallero, M.A., Eggleton, P., Michalak, M., and Murphy-Ullrich, J.E. (2002). The anti-adhesive activity of thrombospondin is mediated by the N-terminal domain of cell surface calreticulin. *J. Biol. Chem.* *277*, 37219–37228.
- Gold, L.I., Rahman, M., Blechman, K.M., Greives, M.R., Churgin, S., Michaels, J., Callaghan, M.J., Cardwell, N.L., Pollins, A.C., Michalak, M., et al. (2006). Overview of the role for calreticulin in the enhancement of wound healing through multiple biological effects. *J. Investig. Dermatol. Symp. Proc. Soc. Investig. Dermatol. Inc Eur. Soc. Dermatol. Res.* *11*, 57–65.
- González, E., Rico, G., Mendoza, G., Ramos, F., García, G., Morán, P., Valadez, A., Melendro, E.I., and Ximénez, C. (2002). Calreticulin-like molecule in trophozoites of *Entamoeba histolytica* HM1:IMSS (Swissprot: accession P83003). *Am. J. Trop. Med. Hyg.* *67*, 636–639.
- González, E., de Leon, M. del C.G., Meza, I., Ocadiz-Delgado, R., Gariglio, P., Silva-Olivares, A., Galindo-Gómez, S., Shibayama, M., Morán, P., Valadez, A., et al. (2011). *Entamoeba histolytica* calreticulin: an endoplasmic reticulum protein expressed by trophozoites into experimentally induced amoebic liver abscesses. *Parasitol. Res.* *108*, 439–449.
- Gopalakrishnapai, J., Gupta, G., Karthikeyan, T., Sinha, S., Kandiah, E., Gemma, E., Oscarson, S., and Suroliya, A. (2006). Isothermal titration calorimetric study defines the substrate binding residues of calreticulin. *Biochem. Biophys. Res. Commun.* *351*, 14–20.
- Görlich, D., and Rapoport, T.A. (1993). Protein translocation into proteoliposomes reconstituted from purified components of the endoplasmic reticulum membrane. *Cell* *75*, 615–630.
- Gregory, L.A., Thielens, N.M., Arlaud, G.J., Fontecilla-Camps, J.C., and Gaboriaud, C. (2003). X-ray structure of the Ca²⁺-binding interaction domain of C1s. Insights into the assembly of the C1 complex of complement. *J. Biol. Chem.* *278*, 32157–32164.
- Groeneveld, T.W.L., Oroszlán, M., Owens, R.T., Faber-Krol, M.C., Bakker, A.C., Arlaud, G.J., McQuillan, D.J., Kishore, U., Daha, M.R., and Roos, A. (2005). Interactions of the extracellular matrix proteoglycans decorin and biglycan with C1q and collectins. *J. Immunol. Baltim. Md* *1950* *175*, 4715–4723.
- Guo, L., Groenendyk, J., Papp, S., Dabrowska, M., Knoblach, B., Kay, C., Parker, J.M.R., Opas, M., and Michalak, M. (2003). Identification of an N-domain histidine essential for chaperone function in calreticulin. *J. Biol. Chem.* *278*, 50645–50653.

Guo, X., Barroso, L., Becker, S.M., Lyerly, D.M., Vedvick, T.S., Reed, S.G., Petri, W.A., and Houpt, E.R. (2009). Protection against intestinal amebiasis by a recombinant vaccine is transferable by T cells and mediated by gamma interferon. *Infect. Immun.* 77, 3909–3918.

H

Haas, I.G., and Wabl, M. (1983). Immunoglobulin heavy chain binding protein. *Nature* 306, 387–389.

Hadders, M.A., Beringer, D.X., and Gros, P. (2007). Structure of C8alpha-MACPF reveals mechanism of membrane attack in complement immune defense. *Science* 317, 1552–1554.

Hammer, G.E., and Shastri, N. (2007). Construction and destruction of MHC class I in the peptide-loading complex. *Nat. Immunol.* 8, 793–794.

Hammond, C., Braakman, I., and Helenius, A. (1994). Role of N-linked oligosaccharide recognition, glucose trimming, and calnexin in glycoprotein folding and quality control. *Proc. Natl. Acad. Sci. U. S. A.* 91, 913–917.

Hanson, D.C., Siegel, R.C., and Schumaker, V.N. (1985). Segmental flexibility of the C1q subcomponent of human complement and its possible role in the immune response. *J. Biol. Chem.* 260, 3576–3583.

Haque, R., Ali, I.M., Sack, R.B., Farr, B.M., Ramakrishnan, G., and Petri, W.A. (2001). Amebiasis and mucosal IgA antibody against the *Entamoeba histolytica* adherence lectin in Bangladeshi children. *J. Infect. Dis.* 183, 1787–1793.

Haque, R., Mondal, D., Duggal, P., Kabir, M., Roy, S., Farr, B.M., Sack, R.B., and Petri, W.A. (2006). *Entamoeba histolytica* infection in children and protection from subsequent amebiasis. *Infect. Immun.* 74, 904–909.

Haque, R., Mondal, D., Shu, J., Roy, S., Kabir, M., Davis, A.N., Duggal, P., and Petri, W.A. (2007). Correlation of interferon-gamma production by peripheral blood mononuclear cells with childhood malnutrition and susceptibility to amebiasis. *Am. J. Trop. Med. Hyg.* 76, 340–344.

Hayashi, T., Rizzuto, R., Hajnoczky, G., and Su, T.-P. (2009). MAM: more than just a housekeeper. *Trends Cell Biol.* 19, 81–88.

Hebert, D.N., and Molinari, M. (2007). In and out of the ER: protein folding, quality control, degradation, and related human diseases. *Physiol. Rev.* 87, 1377–1408.

Hebert, D.N., Foellmer, B., and Helenius, A. (1996). Calnexin and calreticulin promote folding, delay oligomerization and suppress degradation of influenza hemagglutinin in microsomes. *EMBO J.* 15, 2961–2968.

Helenius, A., and Aebi, M. (2004). Roles of N-linked glycans in the endoplasmic reticulum. *Annu. Rev. Biochem.* 73, 1019–1049.

Helk, E., Bernin, H., Ernst, T., Ittrich, H., Jacobs, T., Heeren, J., Tacke, F., Tannich, E., and Lotter, H. (2013). TNF α -mediated liver destruction by Kupffer cells and Ly6Chi monocytes during *Entamoeba histolytica* infection. *PLoS Pathog.* 9, e1003096.

Hendershot, L.M. (2004). The ER function BiP is a master regulator of ER function. *Mt. Sinai J. Med. N. Y.* 71, 289–297.

Hughes-Jones, N.C., and Gardner, B. (1979). Reaction between the isolated globular sub-units of the complement component C1q and IgG-complexes. *Mol. Immunol.* 16, 697–701.

Huston, C.D., Houtp, E.R., Mann, B.J., Hahn, C.S., and Petri, W.A. (2000). Caspase 3-dependent killing of host cells by the parasite *Entamoeba histolytica*. *Cell. Microbiol.* 2, 617–625.

I

Ihara, Y., Cohen-Doyle, M.F., Saito, Y., and Williams, D.B. (1999). Calnexin discriminates between protein conformational states and functions as a molecular chaperone in vitro. *Mol. Cell* 4, 331–341.

Incardona, M.-F., Bourenkov, G.P., Levik, K., Pieritz, R.A., Popov, A.N., and Svensson, O. (2009). EDNA: a framework for plugin-based applications applied to X-ray experiment online data analysis. *J. Synchrotron Radiat.* 16, 872–879.

Ireland, B.S., Brockmeier, U., Howe, C.M., Elliott, T., and Williams, D.B. (2008). Lectin-deficient calreticulin retains full functionality as a chaperone for class I histocompatibility molecules. *Mol. Biol. Cell* 19, 2413–2423.

Isenman, D.E., Kells, D.I., Cooper, N.R., Müller-Eberhard, H.J., and Pangburn, M.K. (1981). Nucleophilic modification of human complement protein C3: correlation of conformational changes with acquisition of C3b-like functional properties. *Biochemistry (Mosc.)* 20, 4458–4467.

J

Jackson, M.R., Cohen-Doyle, M.F., Peterson, P.A., and Williams, D.B. (1994). Regulation of MHC class I transport by the molecular chaperone, calnexin (p88, IP90). *Science* 263, 384–387.

Jaumouillé, V., and Grinstein, S. (2011). Receptor mobility, the cytoskeleton, and particle binding during phagocytosis. *Curr. Opin. Cell Biol.* 23, 22–29.

Jazowiecka-Rakus, J., Jarosz, M., and Szala, S. (2006). Combination of vasostatin gene therapy with cyclophosphamide inhibits growth of B16(F10) melanoma tumours. *Acta Biochim. Pol.* 53, 199–202.

Jeffery, E., Peters, L.R., and Raghavan, M. (2011). The polypeptide binding conformation of calreticulin facilitates its cell-surface expression under conditions of endoplasmic reticulum stress. *J. Biol. Chem.* 286, 2402–2415.

Jiang, H., Cooper, B., Robey, F.A., and Gewurz, H. (1992). DNA binds and activates complement via residues 14–26 of the human C1q A chain. *J. Biol. Chem.* 267, 25597–25601.

Jiang, H., Burdick, D., Glabe, C.G., Cotman, C.W., and Tenner, A.J. (1994). beta-Amyloid activates complement by binding to a specific region of the collagen-like domain of the C1q A chain. *J. Immunol. Baltim. Md 1950* 152, 5050–5059.

Johns, T.G., and Bernard, C.C. (1997). Binding of complement component C1q to myelin oligodendrocyte glycoprotein: a novel mechanism for regulating CNS inflammation. *Mol. Immunol.* *34*, 33–38.

K

Kabani, M., Kelley, S.S., Morrow, M.W., Montgomery, D.L., Sivendran, R., Rose, M.D., Gierasch, L.M., and Brodsky, J.L. (2003). Dependence of endoplasmic reticulum-associated degradation on the peptide binding domain and concentration of BiP. *Mol. Biol. Cell* *14*, 3437–3448.

Kabsch, W. (2010). XDS. *Acta Crystallogr. D Biol. Crystallogr.* *66*, 125–132.

Kantardjieff, K.A., and Rupp, B. (2003). Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals. *Protein Sci. Publ. Protein Soc.* *12*, 1865–1871.

Kaplan, D., Baldi, C., Chiaramonte, M.G., Fernandez, M.M., Levin, M.J., Malchiodi, E., and Baldi, A. (1998). Expression of a recombinant Fab antibody fragment against cruzipain, the major cysteine proteinase of *Trypanosoma cruzi*. *Biochem. Biophys. Res. Commun.* *253*, 53–58.

Kapoor, M., Srinivas, H., Kandiah, E., Gemma, E., Ellgaard, L., Oscarson, S., Helenius, A., and Surolia, A. (2003). Interactions of substrate with calreticulin, an endoplasmic reticulum chaperone. *J. Biol. Chem.* *278*, 6194–6200.

Kapoor, M., Ellgaard, L., Gopalakrishnapai, J., Schirra, C., Gemma, E., Oscarson, S., Helenius, A., and Surolia, A. (2004). Mutational analysis provides molecular insight into the carbohydrate-binding region of calreticulin: pivotal roles of tyrosine-109 and aspartate-135 in carbohydrate recognition. *Biochemistry (Mosc.)* *43*, 97–106.

Kelsall, B.L., and Ravdin, J.I. (1993). Degradation of human IgA by *Entamoeba histolytica*. *J. Infect. Dis.* *168*, 1319–1322.

Khalife, J., Liu, J.L., Pierce, R., Porchet, E., Godin, C., and Capron, A. (1994). Characterization and localization of *Schistosoma mansoni* calreticulin Sm58. *Parasitology* *108* (Pt 5), 527–532.

Kishore, U., and Reid, K.B. (1999). Modular organization of proteins containing C1q-like globular domain. *Immunopharmacology* *42*, 15–21.

Kishore, U., Sontheimer, R.D., Sastry, K.N., Zaner, K.S., Zappi, E.G., Hughes, G.R., Khamashta, M.A., Strong, P., Reid, K.B., and Eggleton, P. (1997). Release of calreticulin from neutrophils may alter C1q-mediated immune functions. *Biochem. J.* *322* (Pt 2), 543–550.

Kishore, U., Kojouharova, M.S., and Reid, K.B.M. (2002). Recent progress in the understanding of the structure-function relationships of the globular head regions of C1q. *Immunobiology* *205*, 355–364.

Kishore, U., Ghai, R., Greenhough, T.J., Shrive, A.K., Bonifati, D.M., Gadjeva, M.G., Waters, P., Kojouharova, M.S., Chakraborty, T., and Agrawal, A. (2004). Structural and functional anatomy of the globular domain of complement protein C1q. *Immunol. Lett.* *95*, 113–128.

Kissoon-Singh, V., Moreau, F., Trusevych, E., and Chadee, K. (2013). *Entamoeba histolytica* exacerbates epithelial tight junction permeability and proinflammatory responses in Muc2(-/-) mice. *Am. J. Pathol.* *182*, 852–865.

Kleshchenko, Y.Y., Moody, T.N., Furtak, V.A., Ochieng, J., Lima, M.F., and Villalta, F. (2004). Human galectin-3 promotes *Trypanosoma cruzi* adhesion to human coronary artery smooth muscle cells. *Infect. Immun.* *72*, 6717–6721.

Knobel, H.R., Villiger, W., and Isliker, H. (1975). Chemical analysis and electron microscopy studies of human C1q prepared by different methods. *Eur. J. Immunol.* *5*, 78–82.

Kojouharova, M.S., Gadjeva, M.G., Tsacheva, I.G., Zlatarova, A., Roumenina, L.T., Tchorbadjieva, M.I., Atanasov, B.P., Waters, P., Urban, B.C., Sim, R.B., et al. (2004). Mutational analyses of the recombinant globular regions of human C1q A, B, and C chains suggest an essential role for arginine and histidine residues in the C1q-IgG interaction. *J. Immunol. Baltim. Md 1950* *172*, 4351–4358.

Kovacs, H., Campbell, I.D., Strong, P., Johnson, S., Ward, F.J., Reid, K.B., and Eggleton, P. (1998). Evidence that C1q binds specifically to CH2-like immunoglobulin gamma motifs present in the autoantigen calreticulin and interferes with complement activation. *Biochemistry (Mosc.)* *37*, 17865–17874.

Kozlov, G., Pocanschi, C.L., Rosenauer, A., Bastos-Aristizabal, S., Gorelik, A., Williams, D.B., and Gehring, K. (2010a). Structural basis of carbohydrate recognition by calreticulin. *J. Biol. Chem.* *285*, 38612–38620.

Kozlov, G., Bastos-Aristizabal, S., Määttänen, P., Rosenauer, A., Zheng, F., Killikelly, A., Trempe, J.-F., Thomas, D.Y., and Gehring, K. (2010b). Structural basis of cyclophilin B binding by the calnexin/calreticulin P-domain. *J. Biol. Chem.* *285*, 35551–35557.

Krautz, G.M., Peterson, J.D., Godsel, L.M., Krettli, A.U., and Engman, D.M. (1998). Human antibody responses to *Trypanosoma cruzi* 70-kD heat-shock proteins. *Am. J. Trop. Med. Hyg.* *58*, 137–143.

Krishnan, V., Xu, Y., Macon, K., Volanakis, J.E., and Narayana, S.V.L. (2009). The structure of C2b, a fragment of complement component C2 produced during C3 convertase formation. *Acta Crystallogr. D Biol. Crystallogr.* *65*, 266–274.

L

Labriola, C., Cazzulo, J.J., and Parodi, A.J. (1999). *Trypanosoma cruzi* calreticulin is a lectin that binds monoglucosylated oligosaccharides but not protein moieties of glycoproteins. *Mol. Biol. Cell* *10*, 1381–1394.

Labriola, C.A., Conte, I.L., López Medus, M., Parodi, A.J., and Caramelo, J.J. (2010). Endoplasmic reticulum calcium regulates the retrotranslocation of *Trypanosoma cruzi* calreticulin to the cytosol. *PLoS One* *5*.

Labriola, C.A., Giraldo, A.M.V., Parodi, A.J., and Caramelo, J.J. (2011). Functional cooperation between BiP and calreticulin in the folding maturation of a glycoprotein in *Trypanosoma cruzi*. *Mol. Biochem. Parasitol.* *175*, 112–117.

- Lachmann, P.J., and Halbwachs, L. (1975). The influence of C3b inactivator (KAF) concentration on the ability of serum to support complement activation. *Clin. Exp. Immunol.* *21*, 109–114.
- Laursen, N.S., Andersen, K.R., Braren, I., Spillner, E., Sottrup-Jensen, L., and Andersen, G.R. (2011). Substrate recognition by complement convertases revealed in the C5-cobra venom factor complex. *EMBO J.* *30*, 606–616.
- Laursen, N.S., Magnani, F., Gottfredsen, R.H., Petersen, S.V., and Andersen, G.R. (2012). Structure, function and control of complement C5 and its proteolytic fragments. *Curr. Mol. Med.* *12*, 1083–1097.
- Leach, M.R., and Williams, D.B. (2004). Lectin-deficient calnexin is capable of binding class I histocompatibility molecules in vivo and preventing their degradation. *J. Biol. Chem.* *279*, 9072–9079.
- Leach, M.R., Cohen-Doyle, M.F., Thomas, D.Y., and Williams, D.B. (2002). Localization of the lectin, ERp57 binding, and polypeptide binding sites of calnexin and calreticulin. *J. Biol. Chem.* *277*, 29686–29697.
- Lejeune, M., Moreau, F., and Chadee, K. (2011). Prostaglandin E2 produced by *Entamoeba histolytica* signals via EP4 receptor and alters claudin-4 to increase ion permeability of tight junctions. *Am. J. Pathol.* *179*, 807–818.
- Li, K., Gor, J., and Perkins, S.J. (2010a). Self-association and domain rearrangements between complement C3 and C3u provide insight into the activation mechanism of C3. *Biochem. J.* *431*, 63–72.
- Li, K., Gor, J., and Perkins, S.J. (2010b). Self-association and domain rearrangements between complement C3 and C3u provide insight into the activation mechanism of C3. *Biochem. J.* *431*, 63–72.
- Li, X., Jiang, L., Wang, Y., Xiao, Y., Huang, Y., Yao, Q., Yang, Y., and Wu, X. (2007). Inhibition of angiogenesis by a novel small peptide consisting of the active fragments of platelet factor-4 and vasostatin. *Cancer Lett.* *256*, 29–32.
- Lièvremon, J.P., Rizzuto, R., Hendershot, L., and Meldolesi, J. (1997). BiP, a major chaperone protein of the endoplasmic reticulum lumen, plays a direct and important role in the storage of the rapidly exchanging pool of Ca²⁺. *J. Biol. Chem.* *272*, 30873–30879.
- Liu, M., Imam, H., Oberg, K., and Zhou, Y. (2005). Gene transfer of vasostatin, a calreticulin fragment, into neuroendocrine tumor cells results in enhanced malignant behavior. *Neuroendocrinology* *82*, 1–10.
- Loftus, B., Anderson, I., Davies, R., Alsmark, U.C.M., Samuelson, J., Amedeo, P., Roncaglia, P., Berriman, M., Hirt, R.P., Mann, B.J., et al. (2005). The genome of the protist parasite *Entamoeba histolytica*. *Nature* *433*, 865–868.
- Lotter, H., González-Roldán, N., Lindner, B., Winau, F., Isibasi, A., Moreno-Lafont, M., Ulmer, A.J., Holst, O., Tannich, E., and Jacobs, T. (2009). Natural killer T cells activated by a lipopeptidophosphoglycan from *Entamoeba histolytica* are critically important to control amebic liver abscess. *PLoS Pathog.* *5*, e1000434.

Lovelace, L.L., Chiswell, B., Slade, D.J., Sodetz, J.M., and Lebioda, L. (2008). Crystal structure of complement protein C8gamma in complex with a peptide containing the C8gamma binding site on C8alpha: implications for C8gamma ligand binding. *Mol. Immunol.* *45*, 750–756.

Lovelace, L.L., Cooper, C.L., Sodetz, J.M., and Lebioda, L. (2011). Structure of human C8 protein provides mechanistic insight into membrane pore formation by complement. *J. Biol. Chem.* *286*, 17585–17592.

M

Ma, L., Luo, L., Qiao, H., Dong, X., Pan, S., Jiang, H., Krissansen, G.W., and Sun, X. (2007). Complete eradication of hepatocellular carcinomas by combined vasostatin gene therapy and B7H3-mediated immunotherapy. *J. Hepatol.* *46*, 98–106.

Machado, F.S., Dutra, W.O., Esper, L., Gollob, K.J., Teixeira, M.M., Factor, S.M., Weiss, L.M., Nagajyothi, F., Tanowitz, H.B., and Garg, N.J. (2012). Current understanding of immunity to *Trypanosoma cruzi* infection and pathogenesis of Chagas disease. *Semin. Immunopathol.* *34*, 753–770.

Malhotra, R., Thiel, S., Reid, K.B., and Sim, R.B. (1990). Human leukocyte C1q receptor binds other soluble proteins with collagen domains. *J. Exp. Med.* *172*, 955–959.

Malhotra, R., Willis, A.C., Jensenius, J.C., Jackson, J., and Sim, R.B. (1993). Structure and homology of human C1q receptor (collectin receptor). *Immunology* *78*, 341–348.

Marie, C., and Petri, W.A. (2014). Regulation of virulence of *Entamoeba histolytica*. *Annu. Rev. Microbiol.* *68*, 493–520.

Marion, S., Laurent, C., and Guillén, N. (2005). Signalization and cytoskeleton activity through myosin IB during the early steps of phagocytosis in *Entamoeba histolytica*: a proteomic approach. *Cell. Microbiol.* *7*, 1504–1518.

Marqués, G., Antón, L.C., Barrio, E., Sánchez, A., Ruiz, S., Gavilanes, F., and Vivanco, F. (1993). Arginine residues of the globular regions of human C1q involved in the interaction with immunoglobulin G. *J. Biol. Chem.* *268*, 10393–10402.

Martens, H.A., Zuurman, M.W., de Lange, A.H.M., Nolte, I.M., van der Steege, G., Navis, G.J., Kallenberg, C.G.M., Seelen, M.A., and Bijl, M. (2009). Analysis of C1q polymorphisms suggests association with systemic lupus erythematosus, serum C1q and CH50 levels and disease severity. *Ann. Rheum. Dis.* *68*, 715–720.

Martin, M., Leffler, J., and Blom, A.M. (2012). Annexin A2 and A5 serve as new ligands for C1q on apoptotic cells. *J. Biol. Chem.* *287*, 33733–33744.

Martin, V., Groenendyk, J., Steiner, S.S., Guo, L., Dabrowska, M., Parker, J.M.R., Müller-Esterl, W., Opas, M., and Michalak, M. (2006). Identification by mutational analysis of amino acid residues essential in the chaperone function of calreticulin. *J. Biol. Chem.* *281*, 2338–2346.

Matlack, K.E., Misselwitz, B., Plath, K., and Rapoport, T.A. (1999). BiP acts as a molecular ratchet during posttranslational transport of prepro-alpha factor across the ER membrane. *Cell* *97*, 553–564.

- Matthews, B.W. (1968). Solvent content of protein crystals. *J. Mol. Biol.* *33*, 491–497.
- McCall, M.N., and Easterbrook-Smith, S.B. (1989). Comparison of the role of tyrosine residues in human IgG and rabbit IgG in binding of complement subcomponent C1q. *Biochem. J.* *257*, 845–851.
- McCoy, J.J., and Mann, B.J. (2005). Proteomic analysis of Gal/GalNAc lectin-associated proteins in *Entamoeba histolytica*. *Exp. Parasitol.* *110*, 220–225.
- McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. *J. Appl. Crystallogr.* *40*, 658–674.
- McGowan, K., Kane, A., Asarkof, N., Wicks, J., Guerina, V., Kellum, J., Baron, S., Gintzler, A.R., and Donowitz, M. (1983). *Entamoeba histolytica* causes intestinal secretion: role of serotonin. *Science* *221*, 762–764.
- McGowan, K., Piver, G., Stoff, J.S., and Donowitz, M. (1990). Role of prostaglandins and calcium in the effects of *Entamoeba histolytica* on colonic electrolyte transport. *Gastroenterology* *98*, 873–880.
- McGreal, E., and Gasque, P. (2002). Structure-function studies of the receptors for complement C1q. *Biochem. Soc. Trans.* *30*, 1010–1014.
- Meirelles, M.N., Juliano, L., Carmona, E., Silva, S.G., Costa, E.M., Murta, A.C., and Scharfstein, J. (1992). Inhibitors of the major cysteinyl proteinase (GP57/51) impair host cell invasion and arrest the intracellular development of *Trypanosoma cruzi* in vitro. *Mol. Biochem. Parasitol.* *52*, 175–184.
- Merle, N.S., Church, S.E., Fremeaux-Bacchi, V., and Roumenina, L.T. (2015a). Complement System Part I - Molecular Mechanisms of Activation and Regulation. *Front. Immunol.* *6*, 262.
- Merle, N.S., Noe, R., Halbwachs-Mecarelli, L., Fremeaux-Bacchi, V., and Roumenina, L.T. (2015b). Complement System Part II: Role in Immunity. *Front. Immunol.* *6*, 257.
- Mery, L., Mesaeli, N., Michalak, M., Opas, M., Lew, D.P., and Krause, K.H. (1996). Overexpression of calreticulin increases intracellular Ca²⁺ storage and decreases store-operated Ca²⁺ influx. *J. Biol. Chem.* *271*, 9332–9339.
- Mevorach, D., Mascarenhas, J.O., Gershov, D., and Elkon, K.B. (1998). Complement-dependent clearance of apoptotic cells by human macrophages. *J. Exp. Med.* *188*, 2313–2320.
- Meyer, D.I., and Dobberstein, B. (1980). Identification and characterization of a membrane component essential for the translocation of nascent proteins across the membrane of the endoplasmic reticulum. *J. Cell Biol.* *87*, 503–508.
- Michalak, M., Robert Parker, J.M., and Opas, M. (2002). Ca²⁺ signaling and calcium binding chaperones of the endoplasmic reticulum. *Cell Calcium* *32*, 269–278.
- Michalak, M., Groenendyk, J., Szabo, E., Gold, L.I., and Opas, M. (2009). Calreticulin, a multi-process calcium-buffering chaperone of the endoplasmic reticulum. *Biochem. J.* *417*, 651–666.
- Min, X., Lemon, B., Tang, J., Liu, Q., Zhang, R., Walker, N., Li, Y., and Wang, Z. (2012). Crystal structure of a single-chain trimer of human adiponectin globular domain. *FEBS Lett.* *586*, 912–917.

Mitra, K., Schaffitzel, C., Shaikh, T., Tama, F., Jenni, S., Brooks, C.L., Ban, N., and Frank, J. (2005). Structure of the E. coli protein-conducting channel bound to a translating ribosome. *Nature* *438*, 318–324.

Molinari, M., Galli, C., Piccaluga, V., Pieren, M., and Paganetti, P. (2002). Sequential assistance of molecular chaperones and transient formation of covalent complexes during protein degradation from the ER. *J. Cell Biol.* *158*, 247–257.

Molinari, M., Eriksson, K.K., Calanca, V., Galli, C., Cresswell, P., Michalak, M., and Helenius, A. (2004). Contrasting functions of calreticulin and calnexin in glycoprotein folding and ER quality control. *Mol. Cell* *13*, 125–135.

Mondal, D., Haque, R., Sack, R.B., Kirkpatrick, B.D., and Petri, W.A. (2009). Attribution of malnutrition to cause-specific diarrheal illness: evidence from a prospective study of preschool children in Mirpur, Dhaka, Bangladesh. *Am. J. Trop. Med. Hyg.* *80*, 824–826.

Mortara, R.A., da Silva, S., Araguth, M.F., Blanco, S.A., and Yoshida, N. (1992). Polymorphism of the 35- and 50-kilodalton surface glycoconjugates of *Trypanosoma cruzi* metacyclic trypomastigotes. *Infect. Immun.* *60*, 4673–4678.

Müller-Loennies, S., MacKenzie, C.R., Patenaude, S.I., Evans, S.V., Kosma, P., Brade, H., Brade, L., and Narang, S. (2000). Characterization of high affinity monoclonal antibodies specific for chlamydial lipopolysaccharide. *Glycobiology* *10*, 121–130.

Müller-Taubenberger, A., Lupas, A.N., Li, H., Ecke, M., Simmeth, E., and Gerisch, G. (2001). Calreticulin and calnexin in the endoplasmic reticulum are important for phagocytosis. *EMBO J.* *20*, 6772–6782.

Muñoz, L.E., Lauber, K., Schiller, M., Manfredi, A.A., and Herrmann, M. (2010a). The role of defective clearance of apoptotic cells in systemic autoimmunity. *Nat. Rev. Rheumatol.* *6*, 280–289.

Muñoz, L.E., Peter, C., Herrmann, M., Wesselborg, S., and Lauber, K. (2010b). Scent of dying cells: the role of attraction signals in the clearance of apoptotic cells and its immunological consequences. *Autoimmun. Rev.* *9*, 425–430.

Murakami, Y., Imamichi, T., and Nagasawa, S. (1993). Characterization of C3a anaphylatoxin receptor on guinea-pig macrophages. *Immunology* *79*, 633–638.

Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* *53*, 240–255.

N

Nakamura, K., Zuppini, A., Arnaudeau, S., Lynch, J., Ahsan, I., Krause, R., Papp, S., De Smedt, H., Parys, J.B., Muller-Esterl, W., et al. (2001). Functional specialization of calreticulin domains. *J. Cell Biol.* *154*, 961–972.

Nanney, L.B., Woodrell, C.D., Greives, M.R., Cardwell, N.L., Pollins, A.C., Bancroft, T.A., Chesser, A., Michalak, M., Rahman, M., Siebert, J.W., et al. (2008). Calreticulin enhances porcine wound repair by diverse biological effects. *Am. J. Pathol.* *173*, 610–630.

- Naresha, S., Suryawanshi, A., Agarwal, M., Singh, B.P., and Joshi, P. (2009). Mapping the complement C1q binding site in *Haemonchus contortus* calreticulin. *Mol. Biochem. Parasitol.* *166*, 42–46.
- Nauta, A.J., Trouw, L.A., Daha, M.R., Tijmsa, O., Nieuwland, R., Schwaeble, W.J., Gingras, A.R., Mantovani, A., Hack, E.C., and Roos, A. (2002). Direct binding of C1q to apoptotic cells and cell blebs induces complement activation. *Eur. J. Immunol.* *32*, 1726–1736.
- Nauta, A.J., Castellano, G., Xu, W., Woltman, A.M., Borrias, M.C., Daha, M.R., van Kooten, C., and Roos, A. (2004). Opsonization with C1q and mannose-binding lectin targets apoptotic cells to dendritic cells. *J. Immunol. Baltim. Md 1950* *173*, 3044–3050.
- Nayal, A., Webb, D.J., and Horwitz, A.F. (2004). Talin: an emerging focal point of adhesion dynamics. *Curr. Opin. Cell Biol.* *16*, 94–98.
- Nde, P.N., Simmons, K.J., Kleshchenko, Y.Y., Pratap, S., Lima, M.F., and Villalta, F. (2006). Silencing of the laminin gamma-1 gene blocks *Trypanosoma cruzi* infection. *Infect. Immun.* *74*, 1643–1648.
- Nde, P.N., Johnson, C.A., Pratap, S., Cardenas, T.C., Kleshchenko, Y.Y., Furtak, V.A., Simmons, K.J., Lima, M.F., and Villalta, F. (2010). Gene network analysis during early infection of human coronary artery smooth muscle cells by *Trypanosoma cruzi* and Its gp83 ligand. *Chem. Biodivers.* *7*, 1051–1064.
- Nde, P.N., Lima, M.F., Johnson, C.A., Pratap, S., and Villalta, F. (2012). Regulation and use of the extracellular matrix by *Trypanosoma cruzi* during early infection. *Front. Immunol.* *3*, 337.
- Ng, F.W., Nguyen, M., Kwan, T., Branton, P.E., Nicholson, D.W., Cromlish, J.A., and Shore, G.C. (1997). p28 Bap31, a Bcl-2/Bcl-XL- and procaspase-8-associated protein in the endoplasmic reticulum. *J. Cell Biol.* *139*, 327–338.
- Nishida, N., Walz, T., and Springer, T.A. (2006). Structural transitions of complement component C3 and its activation products. *Proc. Natl. Acad. Sci. U. S. A.* *103*, 19737–19742.
- 
- Obeid, M., Tesniere, A., Ghiringhelli, F., Fimia, G.M., Apetoh, L., Perfettini, J.-L., Castedo, M., Mignot, G., Panaretakis, T., Casares, N., et al. (2007). Calreticulin exposure dictates the immunogenicity of cancer cell death. *Nat. Med.* *13*, 54–61.
- Ogden, C.A., deCathelineau, A., Hoffmann, P.R., Bratton, D., Ghebrehiwet, B., Fadok, V.A., and Henson, P.M. (2001). C1q and mannose binding lectin engagement of cell surface calreticulin and CD91 initiates macropinocytosis and uptake of apoptotic cells. *J. Exp. Med.* *194*, 781–795.
- Ogle, J.M., and Ramakrishnan, V. (2005). Structural insights into translational fidelity. *Annu. Rev. Biochem.* *74*, 129–177.
- Oliver, J.D., van der Wal, F.J., Bulleid, N.J., and High, S. (1997). Interaction of the thiol-dependent reductase ERp57 with nascent glycoproteins. *Science* *275*, 86–88.

Orr, A.W., Pallero, M.A., and Murphy-Ullrich, J.E. (2002). Thrombospondin stimulates focal adhesion disassembly through Gi- and phosphoinositide 3-kinase-dependent ERK activation. *J. Biol. Chem.* *277*, 20453–20460.

Orr, A.W., Elzie, C.A., Kucik, D.F., and Murphy-Ullrich, J.E. (2003). Thrombospondin signaling through the calreticulin/LDL receptor-related protein co-complex stimulates random and directed cell migration. *J. Cell Sci.* *116*, 2917–2927.

Osorio, L., Ríos, I., Gutiérrez, B., and González, J. (2012). Virulence factors of *Trypanosoma cruzi*: who is who? *Microbes Infect. Inst. Pasteur* *14*, 1390–1402.

Ouaissi, M.A., Afchain, D., Capron, A., and Grimaud, J.A. (1984). Fibronectin receptors on *Trypanosoma cruzi* trypomastigotes and their biological function. *Nature* *308*, 380–382.

Pace, C.N., Vajdos, F., Fee, L., Grimsley, G., and Gray, T. (1995). How to measure and predict the molar absorption coefficient of a protein. *Protein Sci. Publ. Protein Soc.* *4*, 2411–2423.

P

Païdassi, H., Tacnet-Delorme, P., Garlatti, V., Darnault, C., Ghebrehiwet, B., Gaboriaud, C., Arlaud, G.J., and Frchet, P. (2008a). C1q binds phosphatidylserine and likely acts as a multiligand-bridging molecule in apoptotic cell recognition. *J. Immunol. Baltim. Md* *1950* *180*, 2329–2338.

Païdassi, H., Tacnet-Delorme, P., Lunardi, T., Arlaud, G.J., Thielens, N.M., and Frchet, P. (2008b). The lectin-like activity of human C1q and its implication in DNA and apoptotic cell recognition. *FEBS Lett.* *582*, 3111–3116.

Païdassi, H., Tacnet-Delorme, P., Verneret, M., Gaboriaud, C., Houen, G., Duus, K., Ling, W.L., Arlaud, G.J., and Frchet, P. (2011). Investigations on the C1q-calreticulin-phosphatidylserine interactions yield new insights into apoptotic cell recognition. *J. Mol. Biol.* *408*, 277–290.

Pangburn, M.K., Schreiber, R.D., and Müller-Eberhard, H.J. (1981a). Formation of the initial C3 convertase of the alternative complement pathway. Acquisition of C3b-like activities by spontaneous hydrolysis of the putative thioester in native C3. *J. Exp. Med.* *154*, 856–867.

Pangburn, M.K., Schreiber, R.D., and Müller-Eberhard, H.J. (1981b). Formation of the initial C3 convertase of the alternative complement pathway. Acquisition of C3b-like activities by spontaneous hydrolysis of the putative thioester in native C3. *J. Exp. Med.* *154*, 856–867.

Parodi, A.J., Behrens, N.H., Leloir, L.F., and Carminatti, H. (1972). The role of polyprenol-bound saccharides as intermediates in glycoprotein synthesis in liver. *Proc. Natl. Acad. Sci. U. S. A.* *69*, 3268–3272.

Pereira-Chioccola, V.L., Acosta-Serrano, A., Correia de Almeida, I., Ferguson, M.A., Souto-Padron, T., Rodrigues, M.M., Travassos, L.R., and Schenkman, S. (2000). Mucin-like molecules form a negatively charged coat that protects *Trypanosoma cruzi* trypomastigotes from killing by human anti-alpha-galactosyl antibodies. *J. Cell Sci.* *113* (Pt 7), 1299–1307.

Perkins, S.J., Villiers, C.L., Arlaud, G.J., Boyd, J., Burton, D.R., Colomb, M.G., and Dwek, R.A. (1984). Neutron scattering studies of subcomponent C1q of first component C1 of human complement and its association with subunit C1r2C1s2 within C1. *J. Mol. Biol.* *179*, 547–557.

- Perry, A.J., Wijeyewickrema, L.C., Wilmann, P.G., Gunzburg, M.J., D'Andrea, L., Irving, J.A., Pang, S.S., Duncan, R.C., Wilce, J.A., Whisstock, J.C., et al. (2013). A molecular switch governs the interaction between the human complement protease C1s and its substrate, complement C4. *J. Biol. Chem.* *288*, 15821–15829.
- Peters, L.R., and Raghavan, M. (2011). Endoplasmic reticulum calcium depletion impacts chaperone secretion, innate immunity, and phagocytic uptake of cells. *J. Immunol. Baltim. Md 1950* *187*, 919–931.
- Peterson, K.M., Shu, J., Duggal, P., Haque, R., Mondal, D., and Petri, W.A. (2010). Association between TNF-alpha and *Entamoeba histolytica* diarrhea. *Am. J. Trop. Med. Hyg.* *82*, 620–625.
- Petoukhov, M.V., Franke, D., Shkumatov, A.V., Tria, G., Kikhney, A.G., Gajda, M., Gorba, C., Mertens, H.D.T., Konarev, P.V., and Svergun, D.I. (2012). New developments in the ATSAS program package for small-angle scattering data analysis. *J. Appl. Crystallogr.* *45*, 342–350.
- Petri, W.A., Haque, R., and Mann, B.J. (2002). The bittersweet interface of parasite and host: lectin-carbohydrate interactions during human invasion by the parasite *Entamoeba histolytica*. *Annu. Rev. Microbiol.* *56*, 39–64.
- Pflieger, D., Przybylski, C., Gonnet, F., Le Caer, J.-P., Lunardi, T., Arlaud, G.J., and Daniel, R. (2010). Analysis of human C1q by combined bottom-up and top-down mass spectrometry: detailed mapping of post-translational modifications and insights into the C1r/C1s binding sites. *Mol. Cell. Proteomics MCP* *9*, 593–610.
- Pickering, M.C., Botto, M., Taylor, P.R., Lachmann, P.J., and Walport, M.J. (2000). Systemic lupus erythematosus, complement deficiency, and apoptosis. *Adv. Immunol.* *76*, 227–324.
- Pike, S.E., Yao, L., Jones, K.D., Cherney, B., Appella, E., Sakaguchi, K., Nakhasi, H., Teruya-Feldstein, J., Wirth, P., Gupta, G., et al. (1998). Vasostatin, a calreticulin fragment, inhibits angiogenesis and suppresses tumor growth. *J. Exp. Med.* *188*, 2349–2356.
- Pike, S.E., Yao, L., Setsuda, J., Jones, K.D., Cherney, B., Appella, E., Sakaguchi, K., Nakhasi, H., Atreya, C.D., Teruya-Feldstein, J., et al. (1999). Calreticulin and calreticulin fragments are endothelial cell inhibitors that suppress tumor growth. *Blood* *94*, 2461–2468.
- Pocanschi, C.L., Kozlov, G., Brockmeier, U., Brockmeier, A., Williams, D.B., and Gehring, K. (2011). Structural and functional relationships between the lectin and arm domains of calreticulin. *J. Biol. Chem.* *286*, 27266–27277.
- Pollock, S., Kozlov, G., Pelletier, M.-F., Trempe, J.-F., Jansen, G., Sitnikov, D., Bergeron, J.J.M., Gehring, K., Ekiel, I., and Thomas, D.Y. (2004). Specific interaction of ERp57 and calnexin determined by NMR spectroscopy and an ER two-hybrid system. *EMBO J.* *23*, 1020–1029.
- Poon, P.H., Schumaker, V.N., Phillips, M.L., and Strang, C.J. (1983). Conformation and restricted segmental flexibility of C1, the first component of human complement. *J. Mol. Biol.* *168*, 563–577.
- Prathap, K., and Gilman, R. (1970). The histopathology of acute intestinal amebiasis. A rectal biopsy study. *Am. J. Pathol.* *60*, 229–246.
- Preissner, K.T., Podack, E.R., and Müller-Eberhard, H.J. (1985). The membrane attack complex of complement: relation of C7 to the metastable membrane binding site of the intermediate complex C5b-7. *J. Immunol. Baltim. Md 1950* *135*, 445–451.

Puig, A., and Gilbert, H.F. (1994). Anti-chaperone behavior of BiP during the protein disulfide isomerase-catalyzed refolding of reduced denatured lysozyme. *J. Biol. Chem.* *269*, 25889–25896.

Q

Quartier, P., Potter, P.K., Ehrenstein, M.R., Walport, M.J., and Botto, M. (2005). Predominant role of IgM-dependent activation of the classical pathway in the clearance of dying cells by murine bone marrow-derived macrophages in vitro. *Eur. J. Immunol.* *35*, 252–260.

Que, X., and Reed, S.L. (2000). Cysteine proteinases and the pathogenesis of amebiasis. *Clin. Microbiol. Rev.* *13*, 196–206.

Que, X., Kim, S.-H., Sajid, M., Eckmann, L., Dinarello, C.A., McKerrow, J.H., and Reed, S.L. (2003). A surface amebic cysteine proteinase inactivates interleukin-18. *Infect. Immun.* *71*, 1274–1280.

Quiñones, W., Urbina, J.A., Dubourdieu, M., and Luis Concepción, J. (2004). The glycosome membrane of *Trypanosoma cruzi* epimastigotes: protein and lipid composition. *Exp. Parasitol.* *106*, 135–149.

Raghavan, M., Wijeyesakere, S.J., Peters, L.R., and Del Cid, N. (2013). Calreticulin in the immune system: ins and outs. *Trends Immunol.* *34*, 13–21.

R

Ragland, B.D., Ashley, L.S., Vaux, D.L., and Petri, W.A. (1994). *Entamoeba histolytica*: target cells killed by trophozoites undergo DNA fragmentation which is not blocked by Bcl-2. *Exp. Parasitol.* *79*, 460–467.

Rajala, M.W., Patterson, C.M., Opp, J.S., Foltin, S.K., Young, V.B., and Myers, M.G. (2014). Leptin acts independently of food intake to modulate gut microbial composition in male mice. *Endocrinology* *155*, 748–757.

Ralston, K.S., and Petri, W.A. (2011). Tissue destruction and invasion by *Entamoeba histolytica*. *Trends Parasitol.* *27*, 254–263.

Ralston, K.S., Solga, M.D., Mackey-Lawrence, N.M., Somlata, null, Bhattacharya, A., and Petri, W.A. (2014). Trophocytosis by *Entamoeba histolytica* contributes to cell killing and tissue invasion. *Nature* *508*, 526–530.

Ramírez, G., Valck, C., Molina, M.C., Ribeiro, C.H., López, N., Sánchez, G., Ferreira, V.P., Billetta, R., Aguilar, L., Maldonado, I., et al. (2011a). *Trypanosoma cruzi* calreticulin: a novel virulence factor that binds complement C1 on the parasite surface and promotes infectivity. *Immunobiology* *216*, 265–273.

Ramírez, G., Valck, C., Ferreira, V.P., López, N., and Ferreira, A. (2011b). Extracellular *Trypanosoma cruzi* calreticulin in the host-parasite interplay. *Trends Parasitol.* *27*, 115–122.

Ramírez, G., Valck, C., Aguilar, L., Kemmerling, U., López-Muñoz, R., Cabrera, G., Morello, A., Ferreira, J., Maya, J.D., Galanti, N., et al. (2012). Roles of *Trypanosoma cruzi* calreticulin in parasite-host interactions and in tumor growth. *Mol. Immunol.* *52*, 133–140.

- Ramírez-Tolosa, G., Aguilar-Guzmán, L., Valck, C., Abello, P., and Ferreira, A. (2014). Is it all That Bad When Living with an Intracellular Protozoan? The Role of *Trypanosoma cruzi* Calreticulin in Angiogenesis and Tumor Growth. *Front. Oncol.* *4*, 382.
- Ramos, R., Juri, M., Ramos, A., Hoecker, G., Lavandero, S., Pena, P., Morello, A., Repetto, Y., Aguillon, J.C., and Ferreira, A. (1991). An immunogenetically defined and immunodominant *Trypanosoma cruzi* antigen. *Am. J. Trop. Med. Hyg.* *44*, 314–322.
- Ramos-Martínez, E., Olivos-García, A., Saavedra, E., Nequiz, M., Sánchez, E.C., Tello, E., El-Hafidi, M., Saralegui, A., Pineda, E., Delgado, J., et al. (2009). *Entamoeba histolytica*: oxygen resistance and virulence. *Int. J. Parasitol.* *39*, 693–702.
- Ravdin, J.I., and Guerrant, R.L. (1981). Role of adherence in cytopathogenic mechanisms of *Entamoeba histolytica*. Study with mammalian tissue culture cells and human erythrocytes. *J. Clin. Invest.* *68*, 1305–1313.
- Ravdin, J.I., Croft, B.Y., and Guerrant, R.L. (1980). Cytopathogenic mechanisms of *Entamoeba histolytica*. *J. Exp. Med.* *152*, 377–390.
- Ravdin, J.I., Sperelakis, N., and Guerrant, R.L. (1982). Effect of ion channel inhibitors on the cytopathogenicity of *Entamoeba histolytica*. *J. Infect. Dis.* *146*, 335–340.
- Ravdin, J.I., Murphy, C.F., Guerrant, R.L., and Long-Krug, S.A. (1985). Effect of antagonists of calcium and phospholipase A on the cytopathogenicity of *Entamoeba histolytica*. *J. Infect. Dis.* *152*, 542–549.
- Ravdin, J.I., Moreau, F., Sullivan, J.A., Petri, W.A., and Mandell, G.L. (1988). Relationship of free intracellular calcium to the cytolytic activity of *Entamoeba histolytica*. *Infect. Immun.* *56*, 1505–1512.
- Rawal, N., and Pangburn, M.K. (2001). Structure/function of C5 convertases of complement. *Int. Immunopharmacol.* *1*, 415–422.
- Redman, C.M. (1969). Biosynthesis of serum proteins and ferritin by free and attached ribosomes of rat liver. *J. Biol. Chem.* *244*, 4308–4315.
- Reed, S.L., Ember, J.A., Herdman, D.S., DiScipio, R.G., Hugli, T.E., and Gigli, I. (1995). The extracellular neutral cysteine proteinase of *Entamoeba histolytica* degrades anaphylatoxins C3a and C5a. *J. Immunol. Baltim. Md 1950* *155*, 266–274.
- Reid, K.B. (1974). A collagen-like amino acid sequence in a polypeptide chain of human C1q (a subcomponent of the first component of complement). *Biochem. J.* *141*, 189–203.
- Reid, K.B. (1982). C1q. *Methods Enzymol.* *82 Pt A*, 319–324.
- Reid, K.B. (1983). Proteins involved in the activation and control of the two pathways of human complement. *Biochem. Soc. Trans.* *11*, 1–12.
- Reid, K.B., Sim, R.B., and Faiers, A.P. (1977). Inhibition of the reconstitution of the haemolytic activity of the first component of human complement by a pepsin-derived fragment of subcomponent C1q. *Biochem. J.* *161*, 239–245.
- Rhodes, D.C. (2000). Binding of Tamm-Horsfall protein to complement 1q measured by ELISA and resonant mirror biosensor techniques under various ionic-strength conditions. *Immunol. Cell Biol.* *78*, 474–482.

- Ribeiro, C.H., López, N.C., Ramírez, G.A., Valck, C.E., Molina, M.C., Aguilar, L., Rodríguez, M., Maldonado, I., Martínez, R., González, C., et al. (2009). Trypanosoma cruzi calreticulin: a possible role in Chagas' disease autoimmunity. *Mol. Immunol.* *46*, 1092–1099.
- Ricklin, D., Hajishengallis, G., Yang, K., and Lambris, J.D. (2010). Complement: a key system for immune surveillance and homeostasis. *Nat. Immunol.* *11*, 785–797.
- Rico, G., Leandro, E., Rojas, S., Giménez, J.A., and Kretschmer, R.R. (2003). The monocyte locomotion inhibitory factor produced by Entamoeba histolytica inhibits induced nitric oxide production in human leukocytes. *Parasitol. Res.* *90*, 264–267.
- Rimoldi, M.T., Tenner, A.J., Bobak, D.A., and Joiner, K.A. (1988). Enhanced invasion of mononuclear phagocytes by serum-treated Trypanosoma cruzi is due to C1q. *Mem. Inst. Oswaldo Cruz* *83 Suppl 1*, 456–458.
- Rimoldi, M.T., Tenner, A.J., Bobak, D.A., and Joiner, K.A. (1989). Complement component C1q enhances invasion of human mononuclear phagocytes and fibroblasts by Trypanosoma cruzi trypomastigotes. *J. Clin. Invest.* *84*, 1982–1989.
- Rizvi, S.M., Mancino, L., Thammavongsa, V., Cantley, R.L., and Raghavan, M. (2004). A polypeptide binding conformation of calreticulin is induced by heat shock, calcium depletion, or by deletion of the C-terminal acidic region. *Mol. Cell* *15*, 913–923.
- Rodriguez, E., Nan, R., Li, K., Gor, J., and Perkins, S.J. (2014). A revised mechanism for the activation of complement C3 to C3b: a molecular explanation of a disease-associated polymorphism. *J. Biol. Chem.*
- Rossi, V., Bally, I., Thielens, N.M., Esser, A.F., and Arlaud, G.J. (1998). Baculovirus-mediated expression of truncated modular fragments from the catalytic region of human complement serine protease C1s. Evidence for the involvement of both complement control protein modules in the recognition of the C4 protein substrate. *J. Biol. Chem.* *273*, 1232–1239.
- Rossmann, M.G. (2001). Molecular replacement--historical background. *Acta Crystallogr. D Biol. Crystallogr.* *57*, 1360–1366.
- Rouet-Benzineb, P., Aparicio, T., Guilmeau, S., Pouzet, C., Descatoire, V., Buyse, M., and Bado, A. (2004). Leptin counteracts sodium butyrate-induced apoptosis in human colon cancer HT-29 cells via NF-kappaB signaling. *J. Biol. Chem.* *279*, 16495–16502.
- Roumenina, L.T., Kantardjiev, A.A., Atanasov, B.P., Waters, P., Gadjeva, M., Reid, K.B.M., Mantovani, A., Kishore, U., and Kojouharova, M.S. (2005). Role of Ca²⁺ in the electrostatic stability and the functional activity of the globular domain of human C1q. *Biochemistry (Mosc.)* *44*, 14097–14109.
- Roumenina, L.T., Ruseva, M.M., Zlatarova, A., Ghai, R., Kolev, M., Olova, N., Gadjeva, M., Agrawal, A., Bottazzi, B., Mantovani, A., et al. (2006). Interaction of C1q with IgG1, C-reactive protein and pentraxin 3: mutational studies using recombinant globular head modules of human C1q A, B, and C chains. *Biochemistry (Mosc.)* *45*, 4093–4104.
- Roumenina, L.T., Popov, K.T., Bureeva, S.V., Kojouharova, M., Gadjeva, M., Rabheru, S., Thakrar, R., Kaplun, A., and Kishore, U. (2008). Interaction of the globular domain of human C1q with Salmonella typhimurium lipopolysaccharide. *Biochim. Biophys. Acta* *1784*, 1271–1276.

Roumenina, L.T., Sène, D., Radanova, M., Blouin, J., Halbwegs-Mecarelli, L., Dragon-Durey, M.-A., Fridman, W.H., and Fremeaux-Bacchi, V. (2011). Functional complement C1q abnormality leads to impaired immune complexes and apoptotic cell clearance. *J. Immunol. Baltim. Md 1950* *187*, 4369–4373.

Ruiz, R.C., Favoreto, S., Dorta, M.L., Oshiro, M.E., Ferreira, A.T., Manque, P.M., and Yoshida, N. (1998). Infectivity of *Trypanosoma cruzi* strains is associated with differential expression of surface glycoproteins with differential Ca²⁺ signalling activity. *Biochem. J.* *330 (Pt 1)*, 505–511.

Rutkevich, L.A., and Williams, D.B. (2011). Participation of lectin chaperones and thiol oxidoreductases in protein folding within the endoplasmic reticulum. *Curr. Opin. Cell Biol.* *23*, 157–166.

S

Saito, Y., Ihara, Y., Leach, M.R., Cohen-Doyle, M.F., and Williams, D.B. (1999). Calreticulin functions in vitro as a molecular chaperone for both glycosylated and non-glycosylated proteins. *EMBO J.* *18*, 6718–6729.

Sánchez Valdéz, F.J., Pérez Brandán, C., Zago, M.P., Labriola, C., Ferreira, A., and Basombrío, M.Á. (2013). *Trypanosoma cruzi* carrying a monoallelic deletion of the calreticulin (TcCRT) gene are susceptible to complement mediated killing and defective in their metacyclogenesis. *Mol. Immunol.* *53*, 198–205.

Sánchez-Valdéz, F.J., Pérez Brandán, C., Ramírez, G., Uncos, A.D., Zago, M.P., Cimino, R.O., Cardozo, R.M., Marco, J.D., Ferreira, A., and Basombrío, M.Á. (2014). A monoallelic deletion of the TcCRT gene increases the attenuation of a cultured *Trypanosoma cruzi* strain, protecting against an in vivo virulent challenge. *PLoS Negl. Trop. Dis.* *8*, e2696.

Santana, J.M., Grellier, P., Schrével, J., and Teixeira, A.R. (1997). A *Trypanosoma cruzi*-secreted 80 kDa proteinase with specificity for human collagen types I and IV. *Biochem. J.* *325 (Pt 1)*, 129–137.

Schanda, P., and Brutscher, B. (2005). Very fast two-dimensional NMR spectroscopy for real-time investigation of dynamic events in proteins on the time scale of seconds. *J. Am. Chem. Soc.* *127*, 8014–8015.

Scharfstein, J., Schmitz, V., Morandi, V., Capella, M.M., Lima, A.P., Morrot, A., Juliano, L., and Müller-Esterl, W. (2000). Host cell invasion by *Trypanosoma cruzi* is potentiated by activation of bradykinin B(2) receptors. *J. Exp. Med.* *192*, 1289–1300.

Schenkman, S., Ferguson, M.A., Heise, N., de Almeida, M.L., Mortara, R.A., and Yoshida, N. (1993). Mucin-like glycoproteins linked to the membrane by glycosylphosphatidylinositol anchor are the major acceptors of sialic acid in a reaction catalyzed by trans-sialidase in metacyclic forms of *Trypanosoma cruzi*. *Mol. Biochem. Parasitol.* *59*, 293–303.

Schneider, S., and Zacharias, M. (2012). Atomic resolution model of the antibody Fc interaction with the complement C1q component. *Mol. Immunol.* *51*, 66–72.

Schrag, J.D., Bergeron, J.J., Li, Y., Borisova, S., Hahn, M., Thomas, D.Y., and Cygler, M. (2001). The Structure of calnexin, an ER chaperone involved in quality control of protein folding. *Mol. Cell* *8*, 633–644.

- Sellar, G.C., Blake, D.J., and Reid, K.B. (1991). Characterization and organization of the genes encoding the A-, B- and C-chains of human complement subcomponent C1q. The complete derived amino acid sequence of human C1q. *Biochem. J.* 274 (Pt 2), 481–490.
- Seydel, K.B., and Stanley, S.L. (1998). *Entamoeba histolytica* induces host cell death in amebic liver abscess by a non-Fas-dependent, non-tumor necrosis factor alpha-dependent pathway of apoptosis. *Infect. Immun.* 66, 2980–2983.
- Sim, R.B., Moestrup, S.K., Stuart, G.R., Lynch, N.J., Lu, J., Schwaeble, W.J., and Malhotra, R. (1998). Interaction of C1q and the collectins with the potential receptors calreticulin (cC1qR/collectin receptor) and megalin. *Immunobiology* 199, 208–224.
- Simmons, K.J., Nde, P.N., Kleshchenko, Y.Y., Lima, M.F., and Villalta, F. (2006). Stable RNA interference of host thrombospondin-1 blocks *Trypanosoma cruzi* infection. *FEBS Lett.* 580, 2365–2370.
- Slade, D.J., Lovelace, L.L., Chruszcz, M., Minor, W., Lebioda, L., and Sodetz, J.M. (2008). Crystal structure of the MACPF domain of human complement protein C8 alpha in complex with the C8 gamma subunit. *J. Mol. Biol.* 379, 331–342.
- Smith, M.J., and Koch, G.L. (1989). Multiple zones in the sequence of calreticulin (CRP55, calregulin, HACBP), a major calcium binding ER/SR protein. *EMBO J.* 8, 3581–3586.
- Somlata, null, Bhattacharya, S., and Bhattacharya, A. (2011). A C2 domain protein kinase initiates phagocytosis in the protozoan parasite *Entamoeba histolytica*. *Nat. Commun.* 2, 230.
- Sörgjerd, K., Ghafouri, B., Jonsson, B.-H., Kelly, J.W., Blond, S.Y., and Hammarström, P. (2006). Retention of misfolded mutant transthyretin by the chaperone BiP/GRP78 mitigates amyloidogenesis. *J. Mol. Biol.* 356, 469–482.
- Sosoniuk, E., Vallejos, G., Kenawy, H., Gaboriaud, C., Thielens, N., Fujita, T., Schwaeble, W., Ferreira, A., and Valck, C. (2014). *Trypanosoma cruzi* calreticulin inhibits the complement lectin pathway activation by direct interaction with L-Ficolin. *Mol. Immunol.* 60, 80–85.
- Souto-Padrón, T., Labriola, C.A., and de Souza, W. (2004). Immunocytochemical localisation of calreticulin in *Trypanosoma cruzi*. *Histochem. Cell Biol.* 122, 563–569.
- Spiro, R.G., Zhu, Q., Bhoyroo, V., and Söling, H.D. (1996). Definition of the lectin-like properties of the molecular chaperone, calreticulin, and demonstration of its copurification with endomannosidase from rat liver Golgi. *J. Biol. Chem.* 271, 11588–11594.
- Steinø, A., Jørgensen, C.S., Laursen, I., and Houen, G. (2004). Interaction of C1q with the receptor calreticulin requires a conformational change in C1q. *Scand. J. Immunol.* 59, 485–495.
- Stempin, C., Giordanengo, L., Gea, S., and Cerbán, F. (2002). Alternative activation and increase of *Trypanosoma cruzi* survival in murine macrophages stimulated by cruzipain, a parasite antigen. *J. Leukoc. Biol.* 72, 727–734.
- Stevens, F.J., and Argon, Y. (1999). Protein folding in the ER. *Semin. Cell Dev. Biol.* 10, 443–454.
- Stuart, G.R., Lynch, N.J., Lu, J., Geick, A., Moffatt, B.E., Sim, R.B., and Schwaeble, W.J. (1996). Localisation of the C1q binding site within C1q receptor/calreticulin. *FEBS Lett.* 397, 245–249.

Stuart, G.R., Lynch, N.J., Day, A.J., Schwaeble, W.J., and Sim, R.B. (1997). The C1q and collectin binding site within C1q receptor (cell surface calreticulin). *Immunopharmacology* 38, 73–80.

Suhre, K., and Sanejouand, Y.-H. (2004). ElNemo: a normal mode web server for protein movement analysis and the generation of templates for molecular replacement. *Nucleic Acids Res.* 32, W610–W614.

Sukhotnik, I., Vadasz, Z., Coran, A.G., Lurie, M., Shiloni, E., Hatoum, O.A., and Mogilner, J.G. (2006). Effect of leptin on intestinal re-growth following massive small bowel resection in rat. *Pediatr. Surg. Int.* 22, 9–15.

Swindle, J., Ajioka, J., Eisen, H., Sanwal, B., Jacquemot, C., Browder, Z., and Buck, G. (1988). The genomic organization and transcription of the ubiquitin genes of *Trypanosoma cruzi*. *EMBO J.* 7, 1121–1127.

T

Tacnet-Delorme, P., Chevallier, S., and Arlaud, G.J. (2001). Beta-amyloid fibrils activate the C1 complex of complement under physiological conditions: evidence for a binding site for A beta on the C1q globular regions. *J. Immunol. Baltim. Md* 1950 167, 6374–6381.

Takizawa, F., Tsuji, S., and Nagasawa, S. (1996). Enhancement of macrophage phagocytosis upon iC3b deposition on apoptotic cells. *FEBS Lett.* 397, 269–272.

Tanner, J., Weis, J., Fearon, D., Whang, Y., and Kieff, E. (1987). Epstein-Barr virus gp350/220 binding to the B lymphocyte C3d receptor mediates adsorption, capping, and endocytosis. *Cell* 50, 203–213.

Tarr, J.M., Young, P.J., Morse, R., Shaw, D.J., Haigh, R., Petrov, P.G., Johnson, S.J., Winyard, P.G., and Eggleton, P. (2010). A mechanism of release of calreticulin from cells during apoptosis. *J. Mol. Biol.* 401, 799–812.

Taylor, P.R., Carugati, A., Fadok, V.A., Cook, H.T., Andrews, M., Carroll, M.C., Savill, J.S., Henson, P.M., Botto, M., and Walport, M.J. (2000). A hierarchical role for classical pathway complement proteins in the clearance of apoptotic cells in vivo. *J. Exp. Med.* 192, 359–366.

Terrasse, R., Tacnet-Delorme, P., Moriscot, C., Pérard, J., Schoehn, G., Vernet, T., Thielens, N.M., Di Guilmi, A.M., and Frachet, P. (2012). Human and pneumococcal cell surface glyceraldehyde-3-phosphate dehydrogenase (GAPDH) proteins are both ligands of human C1q protein. *J. Biol. Chem.* 287, 42620–42633.

Tesniere, A., Schlemmer, F., Boige, V., Kepp, O., Martins, I., Ghiringhelli, F., Aymeric, L., Michaud, M., Apetoh, L., Barault, L., et al. (2010). Immunogenic death of colon cancer cells treated with oxaliplatin. *Oncogene* 29, 482–491.

Thammavongsa, V., Mancino, L., and Raghavan, M. (2005). Polypeptide substrate recognition by calnexin requires specific conformations of the calnexin protein. *J. Biol. Chem.* 280, 33497–33505.

Thielens, N.M., Illy, C., Bally, I.M., and Arlaud, G.J. (1994). Activation of human complement serine-proteinase C1r is down-regulated by a Ca(2+)-dependent intramolecular control that is released in the C1 complex through a signal transmitted by C1q. *Biochem. J.* 301 (Pt 2), 509–516.

Thielmann, Y., Weiergräber, O.H., Mohrlüder, J., and Willbold, D. (2009). Structural framework of the GABARAP-calreticulin interface--implications for substrate binding to endoplasmic reticulum chaperones. *FEBS J.* *276*, 1140–1152.

Thomson, S.P., and Williams, D.B. (2005). Delineation of the lectin site of the molecular chaperone calreticulin. *Cell Stress Chaperones* *10*, 242–251.

Tissot, B., Daniel, R., and Place, C. (2003). Interaction of the C1 complex of complement with sulfated polysaccharide and DNA probed by single molecule fluorescence microscopy. *Eur. J. Biochem. FEBS* *270*, 4714–4720.

Tissot, B., Gonnet, F., Iborra, A., Berthou, C., Thielens, N., Arlaud, G.J., and Daniel, R. (2005). Mass spectrometry analysis of the oligomeric C1q protein reveals the B chain as the target of trypsin cleavage and interaction with fucoidan. *Biochemistry (Mosc.)* *44*, 2602–2609.

Tjoelker, L.W., Seyfried, C.E., Eddy, R.L., Byers, M.G., Shows, T.B., Calderon, J., Schreiber, R.B., and Gray, P.W. (1994). Human, mouse, and rat calnexin cDNA cloning: identification of potential calcium binding motifs and gene localization to human chromosome 5. *Biochemistry (Mosc.)* *33*, 3229–3236.

Tran, V.Q., Herdman, D.S., Torian, B.E., and Reed, S.L. (1998). The neutral cysteine proteinase of *Entamoeba histolytica* degrades IgG and prevents its binding. *J. Infect. Dis.* *177*, 508–511.

Trombetta, E.S., Fleming, K.G., and Helenius, A. (2001). Quaternary and domain structure of glycoprotein processing glucosidase II. *Biochemistry (Mosc.)* *40*, 10717–10722.

Tsai, S.W., Poon, P.H., and Schumaker, V.N. (1997). Expression and characterization of a 159 amino acid, N-terminal fragment of human complement component C1s. *Mol. Immunol.* *34*, 1273–1280.

Tschopp, J., Villiger, W., Fuchs, H., Kilchherr, E., and Engel, J. (1980). Assembly of subcomponents C1r and C1s of first component of complement: electron microscopic and ultracentrifugal studies. *Proc. Natl. Acad. Sci. U. S. A.* *77*, 7014–7018.

Turner, C.W., Lima, M.F., and Villalta, F. (2002). *Trypanosoma cruzi* uses a 45-kDa mucin for adhesion to mammalian cells. *Biochem. Biophys. Res. Commun.* *290*, 29–34.

V

Vaithilingam, A., Teixeira, J.E., Miller, P.J., Heron, B.T., and Huston, C.D. (2012). *Entamoeba histolytica* cell surface calreticulin binds human c1q and functions in amebic phagocytosis of host cells. *Infect. Immun.* *80*, 2008–2018.

Valck, C., Ramírez, G., López, N., Ribeiro, C.H., Maldonado, I., Sánchez, G., Ferreira, V.P., Schwaeble, W., and Ferreira, A. (2010). Molecular mechanisms involved in the inactivation of the first component of human complement by *Trypanosoma cruzi* calreticulin. *Mol. Immunol.* *47*, 1516–1521.

Vandivier, R.W., Ogden, C.A., Fadok, V.A., Hoffmann, P.R., Brown, K.K., Botto, M., Walport, M.J., Fisher, J.H., Henson, P.M., and Greene, K.E. (2002). Role of surfactant proteins A, D, and C1q in the clearance of apoptotic cells in vivo and in vitro: calreticulin and CD91 as a common collectin receptor complex. *J. Immunol. Baltim. Md 1950* *169*, 3978–3986.

Vassilakos, A., Cohen-Doyle, M.F., Peterson, P.A., Jackson, M.R., and Williams, D.B. (1996). The molecular chaperone calnexin facilitates folding and assembly of class I histocompatibility molecules. *EMBO J.* *15*, 1495–1506.

Venkatraman Girija, U., Gingras, A.R., Marshall, J.E., Panchal, R., Sheikh, M.A., Gál, P., Schwaeble, W.J., Mitchell, D.A., Moody, P.C.E., and Wallis, R. (2013). Structural basis of the C1q/C1s interaction and its central role in assembly of the C1 complex of complement activation. *Proc. Natl. Acad. Sci. U. S. A.* *110*, 13916–13920.

Verneret, M., Tacnet-Delorme, P., Osman, R., Awad, R., Grichine, A., Kleman, J.-P., and Frachet, P. (2014). Relative contribution of c1q and apoptotic cell-surface calreticulin to macrophage phagocytosis. *J. Innate Immun.* *6*, 426–434.

Vieira, O.V., Botelho, R.J., and Grinstein, S. (2002). Phagosome maturation: aging gracefully. *Biochem. J.* *366*, 689–704.

Villalta, F., Madison, M.N., Kleshchenko, Y.Y., Nde, P.N., and Lima, M.F. (2008). Molecular analysis of early host cell infection by *Trypanosoma cruzi*. *Front. Biosci. J. Virtual Libr.* *13*, 3714–3734.

Vines, R.R., Ramakrishnan, G., Rogers, J.B., Lockhart, L.A., Mann, B.J., and Petri, W.A. (1998). Regulation of adherence and virulence by the *Entamoeba histolytica* lectin cytoplasmic domain, which contains a beta2 integrin motif. *Mol. Biol. Cell* *9*, 2069–2079.

W

Wada, I., Rindress, D., Cameron, P.H., Ou, W.J., Doherty, J.J., Louvard, D., Bell, A.W., Dignard, D., Thomas, D.Y., and Bergeron, J.J. (1991). SSR alpha and associated calnexin are major calcium binding proteins of the endoplasmic reticulum membrane. *J. Biol. Chem.* *266*, 19599–19610.

Walter, P., and Blobel, G. (1980). Purification of a membrane-associated protein complex required for protein translocation across the endoplasmic reticulum. *Proc. Natl. Acad. Sci. U. S. A.* *77*, 7112–7116.

Walter, P., and Blobel, G. (1981). Translocation of proteins across the endoplasmic reticulum III. Signal recognition protein (SRP) causes signal sequence-dependent and site-specific arrest of chain elongation that is released by microsomal membranes. *J. Cell Biol.* *91*, 557–561.

Walter, P., and Blobel, G. (1982). Signal recognition particle contains a 7S RNA essential for protein translocation across the endoplasmic reticulum. *Nature* *299*, 691–698.

Wang, W., and Chadee, K. (1995). *Entamoeba histolytica* suppresses gamma interferon-induced macrophage class II major histocompatibility complex Ia molecule and I-A beta mRNA expression by a prostaglandin E2-dependent mechanism. *Infect. Immun.* *63*, 1089–1094.

Wang, W., and Malcolm, B.A. (1999). Two-stage PCR protocol allowing introduction of multiple mutations, deletions and insertions using QuikChange Site-Directed Mutagenesis. *BioTechniques* *26*, 680–682.

Ware, F.E., Vassilakos, A., Peterson, P.A., Jackson, M.R., Lehrman, M.A., and Williams, D.B. (1995). The molecular chaperone calnexin binds Glc1Man9GlcNAc2 oligosaccharide as an initial step in recognizing unfolded glycoproteins. *J. Biol. Chem.* *270*, 4697–4704.

White, F.H. (1961). Regeneration of native secondary and tertiary structures by air oxidation of reduced ribonuclease. *J. Biol. Chem.* 236, 1353–1360.

Wijeyesakere, S.J., Rizvi, S.M., and Raghavan, M. (2013). Glycan-dependent and -independent interactions contribute to cellular substrate recruitment by calreticulin. *J. Biol. Chem.* 288, 35104–35116.

Wines, B.D., and Easterbrook-Smith, S.B. (1990). Carbodiimide crosslinking of human C1q and rabbit IgG. *Mol. Immunol.* 27, 221–226.

Winters, M.S., Spellman, D.S., and Lambris, J.D. (2005). Solvent accessibility of native and hydrolyzed human complement protein 3 analyzed by hydrogen/deuterium exchange and mass spectrometry. *J. Immunol. Baltim. Md 1950* 174, 3469–3474.

Wlodawer, A., and Hodgson, K.O. (1975). Crystallization and crystal data of monellin. *Proc. Natl. Acad. Sci. U. S. A.* 72, 398–399.

X, Y, Z

Xiao, F., Wei, Y., Yang, L., Zhao, X., Tian, L., Ding, Z., Yuan, S., Lou, Y., Liu, F., Wen, Y., et al. (2002). A gene therapy for cancer based on the angiogenesis inhibitor, vasostatin. *Gene Ther.* 9, 1207–1213.

Ximénez, C., González, E., Nieves, M.E., Silva-Olivares, A., Shibayama, M., Galindo-Gómez, S., Escobar-Herrera, J., García de León, M.D.C., Morán, P., Valadez, A., et al. (2014). *Entamoeba histolytica* and *E. dispar* Calreticulin: inhibition of classical complement pathway and differences in the level of expression in amoebic liver abscess. *BioMed Res. Int.* 2014, 127453.

Xu, Y., Oomen, R., and Klein, M.H. (1994). Residue at position 331 in the IgG1 and IgG4 CH2 domains contributes to their differential ability to bind and activate complement. *J. Biol. Chem.* 269, 3469–3474.

Yadav, S., Gupta, S., Selvaraj, C., Doharey, P.K., Verma, A., Singh, S.K., and Saxena, J.K. (2014). In silico and in vitro studies on the protein-protein interactions between *Brugia malayi* immunomodulatory protein calreticulin and human C1q. *PLoS One* 9, e106413.

Yan, L., and Stanley, S.L. (2001). Blockade of caspases inhibits amebic liver abscess formation in a mouse model of disease. *Infect. Immun.* 69, 7911–7914.

Yao, L., Pike, S.E., and Tosato, G. (2002). Laminin binding to the calreticulin fragment vasostatin regulates endothelial cell function. *J. Leukoc. Biol.* 71, 47–53.

Yoshida, N., Mortara, R.A., Araguth, M.F., Gonzalez, J.C., and Russo, M. (1989). Metacyclic neutralizing effect of monoclonal antibody 10D8 directed to the 35- and 50-kilodalton surface glycoconjugates of *Trypanosoma cruzi*. *Infect. Immun.* 57, 1663–1667.

Yu, Q., Yu, R., and Qin, X. (2010). The good and evil of complement activation in HIV-1 infection. *Cell. Mol. Immunol.* 7, 334–340.

Zaferani, A., Vivès, R.R., van der Pol, P., Hakvoort, J.J., Navis, G.J., van Goor, H., Daha, M.R., Lortat-Jacob, H., Seelen, M.A., and van den Born, J. (2011). Identification of tubular heparan sulfate as a

docking platform for the alternative complement component properdin in proteinuric renal disease. *J. Biol. Chem.* 286, 5359–5367.

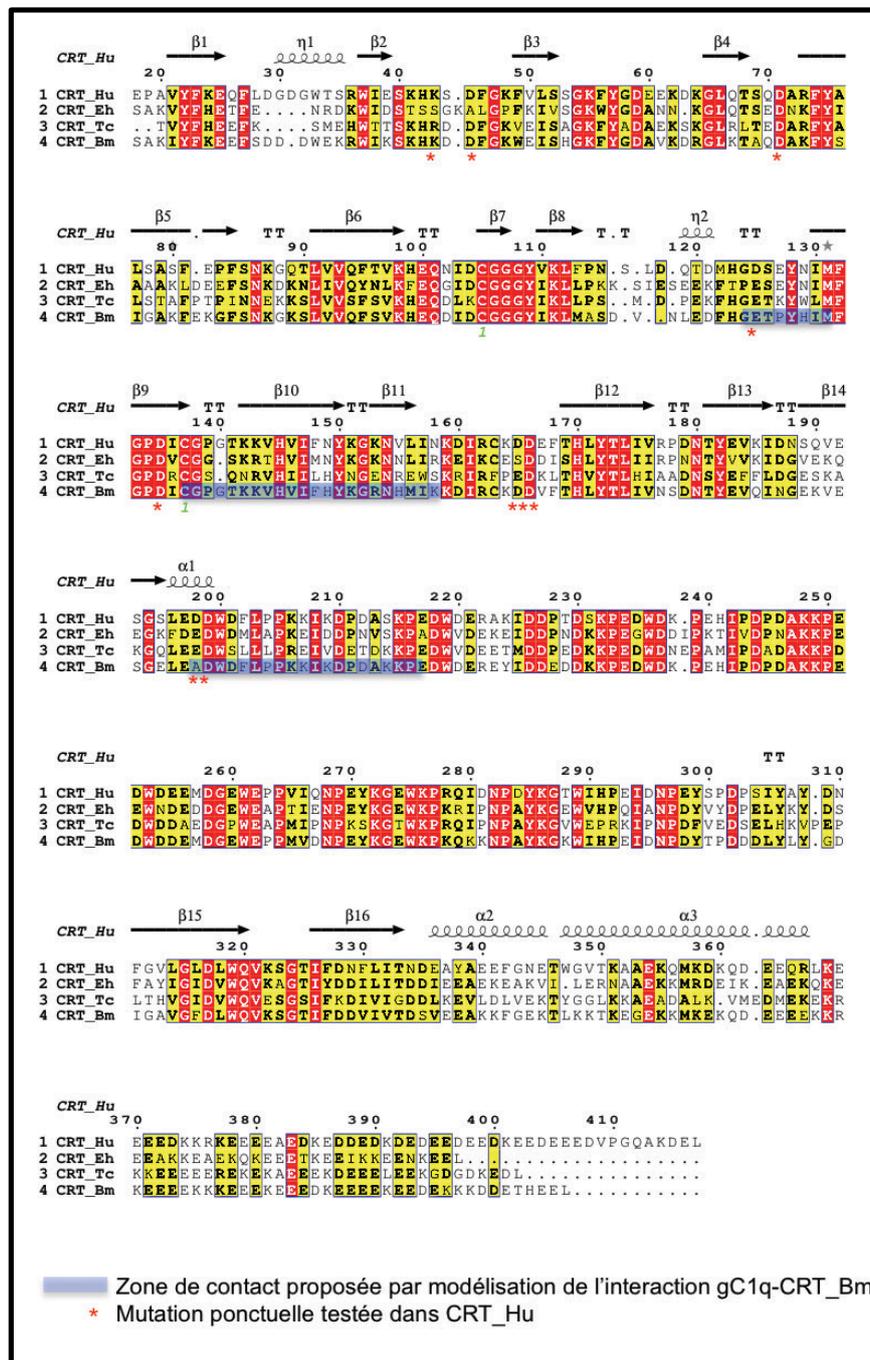
Zapun, A., Jakob, C.A., Thomas, D.Y., and Bergeron, J.J. (1999). Protein folding in a specialized compartment: the endoplasmic reticulum. *Struct. Lond. Engl.* 1993 7, R173–R182.

Zipfel, P.F., Würzner, R., and Skerka, C. (2007). Complement evasion of pathogens: common strategies are shared by diverse organisms. *Mol. Immunol.* 44, 3850–3857.

Zipfel, P.F., Hallström, T., and Riesbeck, K. (2013). Human complement control and complement evasion by pathogenic microbes--tipping the balance. *Mol. Immunol.* 56, 152–160.

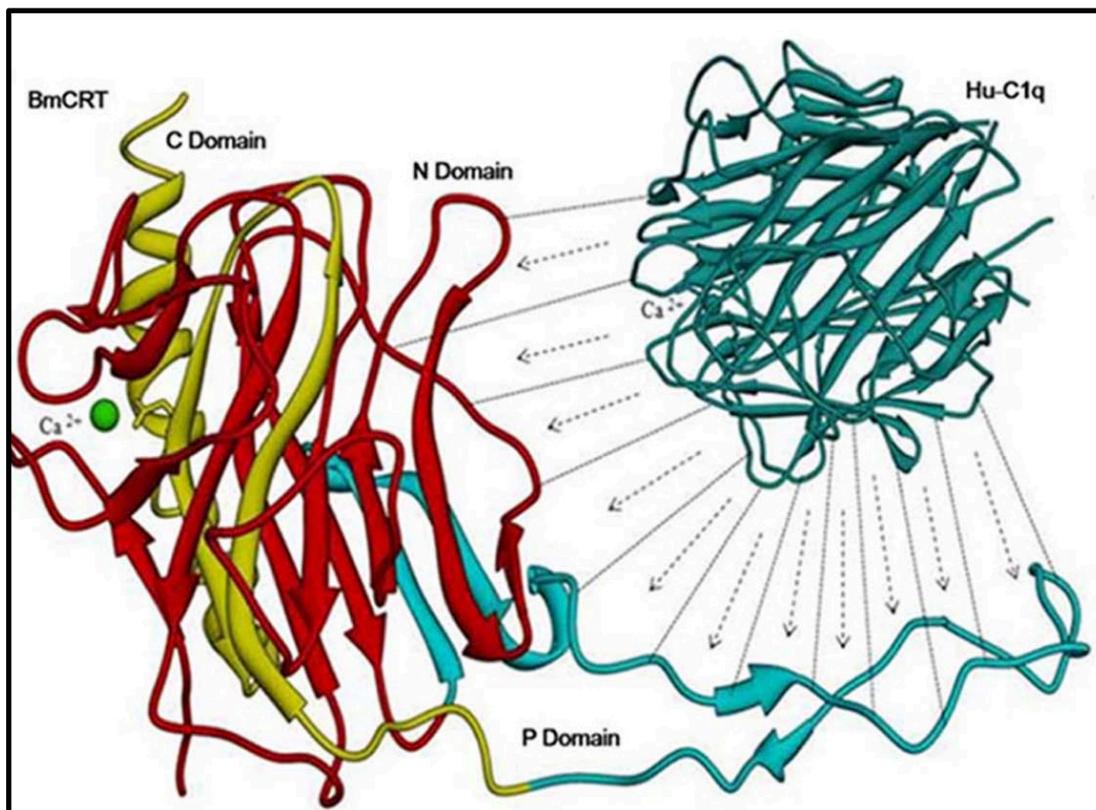
Zlatarova, A.S., Rouseva, M., Roumenina, L.T., Gadjeva, M., Kolev, M., Dobrev, I., Olova, N., Ghai, R., Jensenius, J.C., Reid, K.B.M., et al. (2006). Existence of different but overlapping IgG- and IgM-binding sites on the globular domain of human C1q. *Biochemistry (Mosc.)* 45, 9979–9988.

ANNEXES



ANNEXE 1 : ALIGNEMENT DE SEQUENCE DES CALRETICULINES ETUDIÉES AU COURS DE CETTE THESE ET DE LA CALRETICULINE DE *B. MALAYI*

Les sites de fixation des GRC1q proposés par Yadav et ses collaborateurs sur la CRT de *B. malayi* sont surlignés en bleu. Les mutations déjà effectuées sur la HuCRT au laboratoire sont indiquées par des étoiles rouges. Les structures secondaires sont représentées par des flèches et des ressorts pour les feuillets β et les hélices α respectivement.



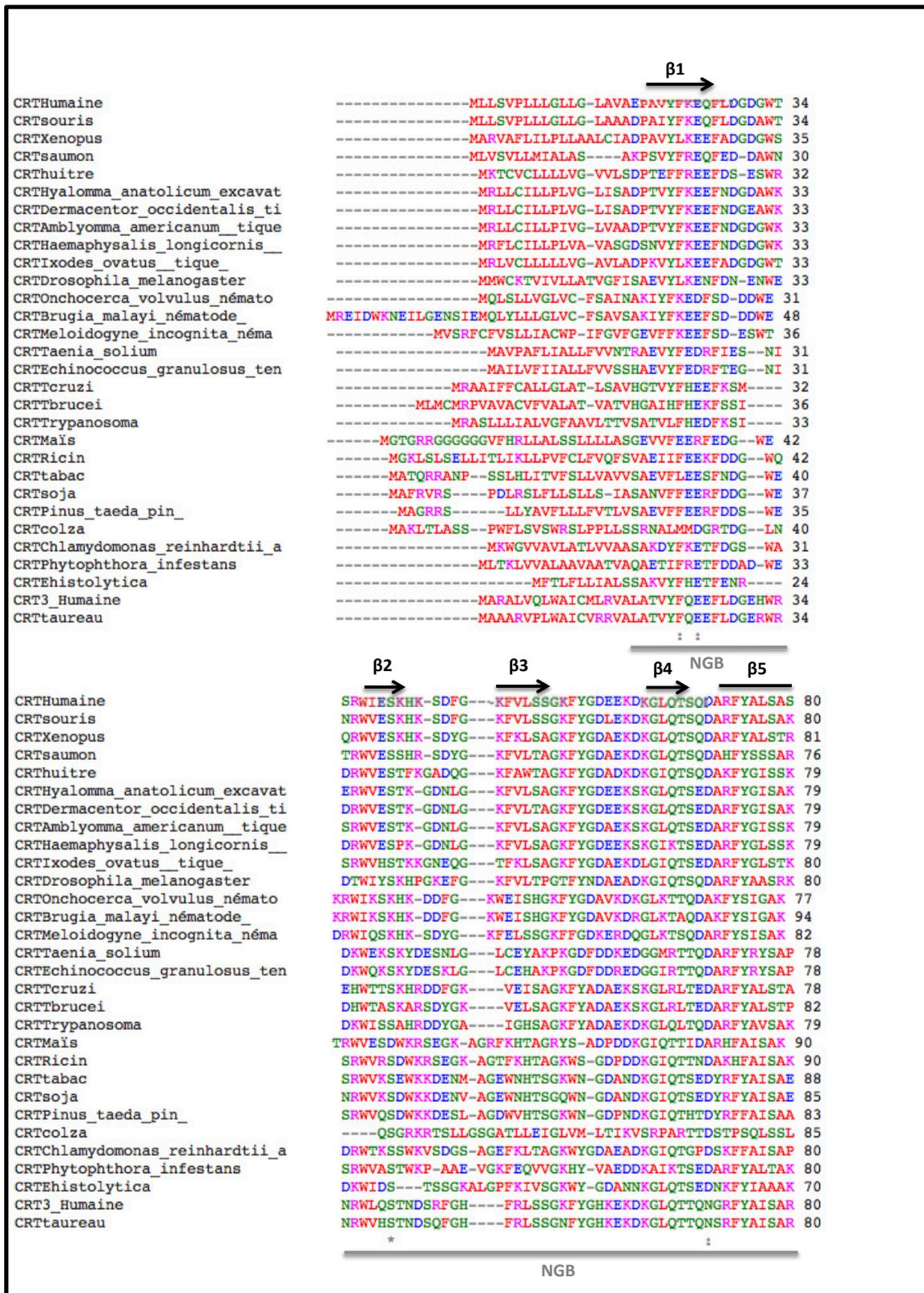
ANNEXE 2 : MODELE D'INTERACTION ENTRE LA CRT DE *B. MALAYI* ET LES GRC1Q OBTENU PAR MODELISATION MOLECULAIRE (YADAV ET AL., 2014)

| | | | | | | | |
|-------------|------------|-------------|------------|-------------|----------------|------------|-----|
| | 10 | 20 | 30 | 40 | 50 | 60 | |
| | DPGNIKDQPR | PAFSAIRRN | PMGGNVVIFD | TVITNQEEPY | QNHSGRFVCT | VPGYYYFTFQ | |
| | 70 | 80 | 90 | 100 | 110 | 120 | |
| | VLSQWEICLS | IVSSSRGQVR | RSLGFCDTTN | KGLFQVVSGG | MVLQLQQGDQ | VWVEKDPKKG | |
| | 130 | 140 | 150 | 160 | 170 | 180 | |
| | HIYQGSEADS | VFSGFLIFPS | AGSGKQKQFS | VFTVTRQTHQ | PPAPNSLIRF | NAVLTNPQGD | |
| | 190 | 200 | 210 | 220 | 230 | 240 | |
| | YDTSTGKFTC | KVPGLYYFVY | HASHTANLCV | LLYRSGVKV | TFCGHTSKTN | QVNSGGVLLR | |
| | 250 | 260 | 270 | 280 | 290 | 300 | |
| | LQVGEEVWLA | VNDYYDMVGI | QGSDSVFSGF | LLFPDGSACA | TQKIAFSATR | TINVPLRRDQ | |
| | 310 | 320 | 330 | 340 | 350 | 360 | |
| | TIRFDHVITN | MNNNYEPRSG | KFTCKVPGLY | YFTYHASSRG | NLCVNLMRGR | ERAQKVVTFC | |
| | 370 | 380 | 390 | 400 | 410 | | |
| | DYAYNTFQVT | TGGMVLKLEQ | GENVFLQATD | KNSLLGMEGA | NSIFSGFLLF | PDMEA | |
| GRC1qr | DPGNIKDQPR | PAFSAIRRN | PMGGNVVIFD | TVITNQEEPY | QNHSGRFVCT | VPGYYYFTFQ | 60 |
| GRC1qplasm. | DPGNIKDQPR | PAFSAIRRN | PMGGNVVIFD | TVITNQEEPY | QNHSGRFVCT | VPGYYYFTFQ | 60 |
| | ***** | | | | | | |
| GRC1qr | VLSQWEICLS | IVSSSRGQVR | RSLGFCDTTN | KGLFQVVSGG | MVLQLQQGDQ | VWVEKDPKKG | 120 |
| GRC1qplasm. | VLSQWEICLS | IVSSSRGQVR | RSLGFCDTTN | KGLFQVVSGG | MVLQLQQGDQ | VWVEKDPKKG | 120 |
| | ***** | | | | | | |
| GRC1qr | HIYQGSEADS | VFSGFLIFPS | AGSGKQKQFS | VFTVTRQTHQ | PPAPNSLIRF | NAVLTNPQGD | 180 |
| GRC1qplasm. | HIYQSEADS | VFSGFLIFPSA | ----- | KFQSVFTVTR | QTHQPPAPNSLIRF | NAVLTNPQGD | 175 |
| | ***** | | | | | | |
| GRC1qr | YDTSTGKFTC | KVPGLYYFVY | HASHTANLCV | LLYRSGVKV | TFCGHTSKTN | QVNSGGVLLR | 240 |
| GRC1qplasm. | YDTSTGKFTC | KVPGLYYFVY | HASHTANLCV | LLYRSGVKV | TFCGHTSKTN | QVNSGGVLLR | 235 |
| | ***** | | | | | | |
| GRC1qr | LQVGEEVWLA | VNDYYDMVGI | QGSDSVFSGF | LLFPDGSACA | TQKIAFSATR | TINVPLRRDQ | 300 |
| GRC1qplasm. | LQVGEEVWLA | VNDYYDMVGI | QGSDSVFSGF | LLFPD----- | TQKIAFSATR | TINVPLRRDQ | 290 |
| | ***** | | | | | | |
| GRC1qr | TIRFDHVITN | MNNNYEPRSG | KFTCKVPGLY | YFTYHASSRG | NLCVNLMRGR | ERAQKVVTFC | 360 |
| GRC1qplasm. | TIRFDHVITN | MNNNYEPRSG | KFTCKVPGLY | YFTYHASSRG | NLCVNLMRGR | ERAQKVVTFC | 350 |
| | ***** | | | | | | |
| GRC1qr | DYAYNTFQVT | TGGMVLKLEQ | GENVFLQATD | KNSLLGMEGAN | SIFSGFLLPD | ME | 415 |
| GRC1qplasm. | DYAYNTFQVT | TGGMVLKLEQ | GENVFLQATD | KNSLLGMEGAN | SIFSGFLLPD | ME | 405 |
| | ***** | | | | | | |

ANNEXE 3 : SEQUENCES DES GRC1Q RECOMBINANTES PRODUITES EN CELLULES HEK 273-F ET ALIGNEMENT DE SEQUENCES ENTRE LES GRC1Q RECOMBINANTES ET LES GRC1Q PLASMATIQUES

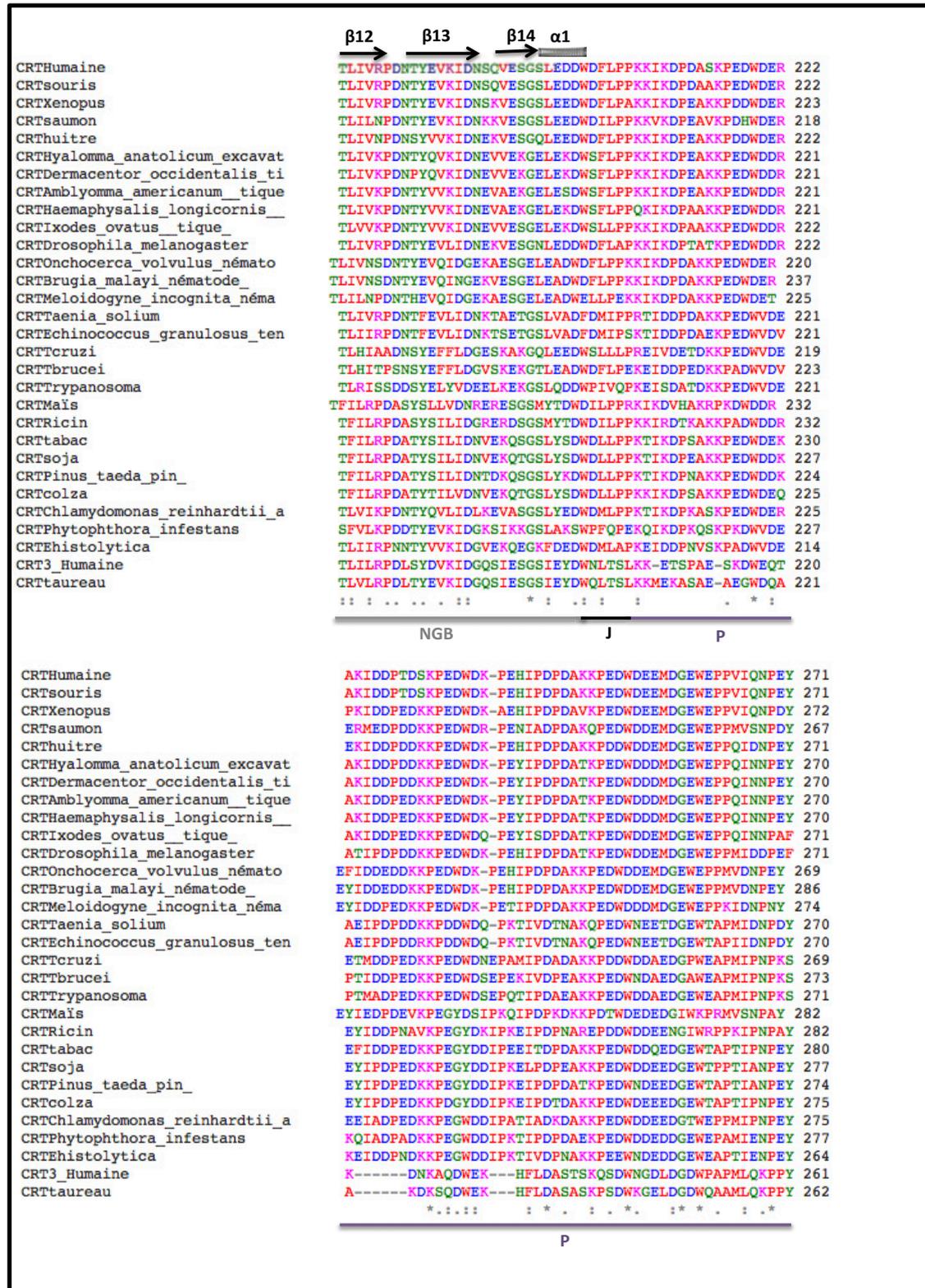
| Protéines | Masse moléculaire (Da) | pI | Absorbance 1% (280nm) |
|--------------------|-----------------------------------|-----------|----------------------------------|
| TcCRTFull | 44255 | 4.88 | 17.07 |
| TcCRTΔP | 37648 | 5.10 | 14.22 |
| TcCRTΔP288 | 34076 | 5.32 | 15.71 |
| TcCRTΔP281 | 33210 | 5.50 | 16.12 |
| TcCRTΔP278 | 32783 | 5.50 | 16.33 |
| TcCRTΔP276 | 32582 | 5.56 | 16.43 |
| TcCRTΔP274 | 32642 | 5.56 | 16.4 |
| TcCRTΔP270 | 32084 | 5.42 | 16.68 |
| TcCRTΔP268 | 31864 | 5.42 | 16.33 |
| TcCRTΔPterΔ288 | 31112 | 5.55 | 13.19 |
| TcCRTNGB | 30119 | 5.77 | 13.62 |
| EhCRTFull | 41026 | 4.83 | 19.61 |
| EhCRTΔP | 31425 | 5.10 | 14.15 |
| EhCRTNGB | 30442 | 5.37 | 14.61 |
| HuCRTFull | 49136 | 4.38 | 16.58 |
| HuCRTNGB | 33329 | 5.32 | 13.21 |
| C1q | 459300 | 8.84 | 6.8 |
| CLR | 189900 | 7.51 | 2.1 (275 nm) |
| GRC1q plasmatique | 48000 | 8.36 | 7 |
| GRC1q recombinante | 45652 | 8.94 | 9.6 |
| gC1qR | 29129 (monomère) | 4.77 | 5.6 |

ANNEXE 4: PARAMETRES PHYSICOCHIMIQUES DES PROTEINES UTILISEES AU COURS DE LA THESE



ANNEXE 5 : ALIGNEMENT DE SEQUENCE DE CALRETICULINES

Alignement de séquence de calréticulines provenant de différentes espèces eucaryotes. Les domaines de la protéine, ainsi que les structures secondaires sont indiqués. Les résidus strictement conservés sont noté du symbole *.



ANNEXE 5 : ALIGNEMENT DE SEQUENCE DE CALRETICULINES

Alignement de séquence de calréticulines provenant de différentes espèces eucaryotes. Les domaines de la protéine, ainsi que les structures secondaires sont indiqués. Les résidus strictement conservés sont noté du symbole *.

| | | | |
|--------------------------------|-------------------------|-----------------------------|--------------------------|
| | | β15 | |
| | | → | |
| CRTHumaine | KGEWKPRQIDNPDKGTWIHP | PEIDNPEYSPDPSIYA-YDNF | GVGLGLDLWQ 320 |
| CRTsouris | KGEWKPRQIDNPDKGTWIHP | PEIDNPEYSPDANIYA-YDSF | AVLGLDLWQ 320 |
| CRTXenopus | KGEWKPRQIDNPDKGTWIHP | PEIDNPEYTPDSTLYQ-YDSF | GVGLGLDLWQ 321 |
| CRTsaumon | KGEWKPRQIDNPDKGTWIHP | PEIDNPEYADSEIYR-FDSI | GVGLGLDLWQ 316 |
| CRTThuitre | KGEWKPKQIDNPDKGTWVHP | PEIDNPEYTPDSNLYK-YDDI | GAVGFDLWQ 320 |
| CRTHyalomma_anatolicum_excavat | KGEWKPKQIDNPAYKGAWVHP | PEIDNPEYADPNLYH-YKEI | CKVGFDFLWQ 319 |
| CRTDermacentor_occidentalis_ti | KGEWKPKQIDNPAYKGAWVHP | PEIDNPEYADPNLYH-YPEI | CKVGFDFLWQ 319 |
| CRTAmblyomma_americanum_tique | KGEWKPKQIDNPAYKGAWVHP | PEIDNPEYADPNLYH-FPEL | CTVGFDFLWQ 319 |
| CRTHaemaphysalis_longicornis__ | KGEWKPKQIDNPAYKGAWVHP | PEIDNPEYTPDKLHA-YKEI | CTVGFDFLWQ 319 |
| CRTIxodes_ovatus_tique_ | KGEWKPKQIDNPAYKGAWVHP | PEIDNPEYTPDKLHS-YKEI | CTVGFDFLWQ 320 |
| CRTDrosophila_melanogaster | KGEWQPKQLDNPNYKGAWHEP | IEANPEYVPDDLYL-RKEI | CTVGFDFLWQ 320 |
| CRTOnchocerca_volvulus_nemato | KGEWKPKQKKNPAYKGKWIHP | PEIEIPDYTPDDNLYV-YDDI | GAI GFDFLWQ 318 |
| CRTBrugia_malayi_nemato | KGEWKPKQKKNPAYKGKWIHP | PEIDNPDYTPDDLYL-YGDI | GAVGFDLWQ 335 |
| CRTMeloidogyne_incognita_nema | KGEWKPKQIKNPNYKGKWIHP | PEIDNPDKVDDLYM-REDW | GSVGI DWQ 323 |
| CRTaenia_solium | KGEWHPKMIPNPAYRGEWKPPQ | IPNPDPHFDESELYA--RTF | AYIGL DLWQ 318 |
| CRTechinococcus_granulosus_ten | KGEWSPRRIIPNPAYKGQWKPPQ | IPNPDPYFDEDELYA--RTF | AYIGL DLWQ 318 |
| CRTTcruzi | KGTWPKRQIPNPAYKGWVPRK | IPNPDPFVDESLELHKVPE | PLTHVGI DWQ 319 |
| CRTBrucei | KGPWAPRKIPNPAYKGPWAPRR | IPNPAYKNDEELYKIPEPL | THVGI DWQ 323 |
| CRTTrypanosoma | KGPWHPKQIPNPAYKGPWEARM | IPNPDPYKPIPDLYKIPEAL | EVGIDVWQ 321 |
| CRTMaïs | KGPWKRRKIKNPNYKGKWKTP | WIDNPEFEDDPLYV-LKPK | LYIGIEVWQ 331 |
| CRTRicin | KGPWRPKRIKNPNYKGKWKVP | YIDNPEFEDDPLYV-LKPK | LYVGVVWQ 331 |
| CRTtabac | KGPWKPKIKNPNYKGKWKAPL | IDNPDFKDDPLYV-FPNL | KYVGI ELWQ 329 |
| CRTsoja | KGPWKPKIKNPNYKGKWKAPL | IDNPDFKDDPLYV-FPNL | KYVGI ELWQ 326 |
| CRTPinus_taeda_pin_ | KGPWKPKIKNPNYKGKWKAPM | IDNPDFKDDPELYV-FPNL | KYVGI ELWQ 323 |
| CRTcolza | MGEWKPKQIKNPNYKGKWEAP | IDNPDFKDDPELYV-FPKL | KYVGI ELWQ 324 |
| CRTChlamydomonas_reinhardtii_a | KGEWKAAMIKNPAYKGIWVAP | IDNPDYVHDDKLYN-FKDL | KVGFELWQ 324 |
| CRTPhytophthora_infestans | KGEWKPKMIDNPDKGEWHP | PLIDNPYFDEAMHNVAKNV | GAI GFELWQ 327 |
| CRTehistolytica | KGEWKPKRIIPNPAYKGWVHP | QIANPDYVDPPELYK-YDSF | AYIGL DVWQ 313 |
| CRT3_Humaine | QDGLKPEGIH---- | KDVVLRHKMNTDYLTYDLSE-FENI | GAI GLELWQ 306 |
| CRTtaureau | QDGLKPEGID---- | KDVVLRHQKMKNS-YLTYDLSE-FENI | GAVGLELWQ 306 |
| | * | * | * |
| | P | J | CGB |
| | ← | ← | ← |
| | β16 | α2 | |
| | → | (.....) | |
| CRTHumaine | VKSGTIFDNFLITNDEAYAE | EFGNETWGVTKAAEKQMK | DKQDEEQLKEE 370 |
| CRTsouris | VKSGTIFDNFLITNDEAYAE | EFGNETWGVTKAAEKQMK | DKQDEEQLKEE 370 |
| CRTXenopus | VKSGTIFDNFLITDDEKYAE | EYGNETWGVTKAAEKMK | EQDEEDRKKQE 371 |
| CRTsaumon | VKSGTIFDNFLITDDATLAE | EVGNETWGVTKDPEKMK | ESQEEERKKLE 366 |
| CRTThuitre | VKSGTIFDNILVTDDEKHA | EEVGNATWGTKDEKMK | EDQDEEERKKRE 370 |
| CRTHyalomma_anatolicum_excavat | VKSGTIFDNLLITDDEEYAR | VHGEEETWAALKDAEK | MKQEEEEAKSK 369 |
| CRTDermacentor_occidentalis_ti | VKSGTIFDNLLITDDEEYAR | VHGEEETWAALKDAEK | MKQEEEEAKSK 369 |
| CRTAmblyomma_americanum_tique | VKSGTIFDNLLITDDEEYAR | VHGEEETWAALKDEEK | MKQEEEEAKSK 369 |
| CRTHaemaphysalis_longicornis__ | VKSGTIFDNLLITDDEEYAR | VHGEEETWGAIKDAEK | MKQEEEEAKNK 369 |
| CRTIxodes_ovatus_tique_ | VKSGTIFDNILVTDDEEYAR | QYGEETWGVTKDGEK | MKQDEKEDKAE 370 |
| CRTDrosophila_melanogaster | VKSGTIFDNVLIITDDELAA | KAAAEVKN-TQAGEK | MKQAEDEVQRKDE 369 |
| CRTOnchocerca_volvulus_nemato | VKSGTIFDDVIVTDSVEEA | KKFGKTLKIREGEEK | KKK----- 358 |
| CRTBrugia_malayi_nemato | VKSGTIFDDVIVTDSVEEA | KKFGKTLKTKEGEK | MKQDE----- 378 |
| CRTMeloidogyne_incognita_nema | VKSGTIFDNIIVTDSIDEA | KAAHAKETFEPLRDAEK | QKAADEEERKKFE 373 |
| CRTaenia_solium | VKSGTIFDNFIVSDNVSEC | QAHAAYWQKRF | AFEEEEKSFEE-----K 362 |
| CRTechinococcus_granulosus_ten | VKSGTIFDNFIVSDDVSEC | QAHAAYWQKRF | TEEEQEKGFEE-----K 362 |
| CRTTcruzi | VESGSIFKDIVIGDDLK | EVLDLVKTGGKKA | EADALKVMED-----M 363 |
| CRTBrucei | VESGSIFKDIIGDDVKE | VLDIVKSTYDGMKKA | EADALAAFE-----M 365 |
| CRTTrypanosoma | VESGSIFNIIIGDDVKE | VLEIVKGTYGASKKA | ESDAFEAFK-----A 365 |
| CRTMaïs | VKAGSVFDNIIICDDPEY | ARKVVEETWGANR | EAEKEAFEEAK-----E 375 |
| CRTRicin | VKAGSIFDNVLIICDDPEY | AKVIVQVF-ANREI | EKEAFEEAK-----I 374 |
| CRTtabac | VKSGTLFDNVIICDDPEY | AKAIAEETWGRQK | DAEKAFAEEAK-----K 373 |
| CRTsoja | VKSGTLFDNVLITDDPEY | AKQLVEETWGRK | HKDAEKAFAEEAK-----K 370 |
| CRTPinus_taeda_pin_ | VKSGTLFDNLIISDDPEY | AKLAEEETWAKHK | DAEKEAFDEAK-----K 367 |
| CRTcolza | VKSGSLFDNVIICDDPDY | AKLAEEETWGLK | DAEKAFAEIEK-----K 368 |
| CRTChlamydomonas_reinhardtii_a | VKSGSIFDNILVTDDLEA | AKKFAEDTWGRK | HKDEEKAMPDKVKK-----E 368 |
| CRTPhytophthora_infestans | VKSGTLFDNIIITDSEDD | FNAHEEAVLAKT---E | MAAKKQ-----I 367 |
| CRTehistolytica | VKAGTIYDDILITDDIEA | EAEKAKVILERN-A | EKKMRDEIKE-----A 356 |
| CRT3_Humaine | VRSGTIFDNFLITDDEEY | ADNFGKATWGETK | GPPEMDAIQAK----- 349 |
| CRTtaureau | VRSGTIFDNFLITDDEEY | ADNFGKATWGETK | GPPEMDAIQAK----- 349 |
| | * | * | * |
| | CGB | Cter | |

ANNEXE 5 : ALIGNEMENT DE SEQUENCE DE CALRETICULINES

Alignement de séquence de calréticulines provenant de différentes espèces eucaryotes. Les domaines de la protéine, ainsi que les structures secondaires sont indiqués. Les résidus strictement conservés sont noté du symbole *.

| | | |
|---------------------------------------|--|-------|
| CRTHumaine | EEDKKRKEE-----EEADKEDDEDKDEDEEDEDKKE--DEEEDVPGQ | 412 |
| CRTsouris | EEDKKRKEE-----EEADKEDDDDRDEDEEDEDKKE--DEEE-SPGQ | 411 |
| CRTXenopus | EEKKRKEE-----EPQEEEEDDDD--EEEEKEEKEE--EEDEEETP | 411 |
| CRTsaumon | AEMARKEETKDEPGESEEEEEEELEHEEEEEDEEGETGAQEEEEESDS | 416 |
| CRTthuitre | EEKKRKEE-----EDKNKKDDDDDDDEEGGDDDDDE---KKDDEDK | 409 |
| CRTHyalomma_anatolicum_excavat | KEDDAKDED-----EFDEEDKEDKDEDKDEEETTPAP--DEDDH | 408 |
| CRTDermacentor_occidentalis_ti | KEDDAKDED-----EFDEDEDKEDKDEK---EETTTPAP--DDEDH | 405 |
| CRTAmblyomma_americanum_tique | KEDEAKDED-----EFDEEKEEDKEKE--DEETTAAP--EDDDH | 405 |
| CRTHaemaphysalis_longicornis__ | KEDDAKDED-----EFDDDEDK--KEEE--EATTPAPSEDDDH | 405 |
| CRTIxodes_ovatus_tique_ | KKDKEDDKD-----EDFDEEDEDKKEE--EEAPADKETEHEHQH | 408 |
| CRTDrosophila_melanogaster | EEAKKASDK-----DDEDEDDDEEKDDESKQ-----DKDQS | 401 |
| CRTOnchocerca_volvulus_nemato | TKKQKKKEK-----NEKIKKE-----KMKRKRANRKKKK-- | 388 |
| CRTBrugia_malayi_nematode_ | EEKKRKEE-----EEKKKEEKEEEDKKEEKEEKEDEKDDDE | 416 |
| CRTMeloidogyne_incognita_nema | EEKKRKEE-----EESKKKDE-DKDGDEEEDKDEEEKVVVE | 410 |
| CRTTaenia_solium | DKESSTT-----DNVPSDEEDEVLDLEGGDAS----- | 390 |
| CRTEchinococcus_granulosus_ten | EKESSTI-----ESLPEESGDEEVDLEEAGDAS----- | 390 |
| CRTTcruzi | EKEKRRK-----EEEEEREKEKAE--EEKDEEELLEEKGDG | 396 |
| CRTTbrucei | KKKEQDK-----KEENKKE TDGEE--DKKKKE----- | 390 |
| CRTTrypanosoma | EDAAAKA-----EEANTATVETIDNGDTKDAE----- | 392 |
| CRTMaïs | RKAREDR-----EGQKAKDDGGRHR---IHRRHKKTLPQR | 406 |
| CRTRicin | RKAREEE-----EAQRAREEGERRRRERGGYDRRHDRDRYD | 409 |
| CRTtabac | REEEESK-----AAPADSDAEE--DDDADDAD---DAD | 402 |
| CRTsoja | REEEESK-----DDPVDSDADE--EEDADEAGNDSDAE | 402 |
| CRTPinus_taeda_pin_ | KEEEEKKE-----SDEEDTDEKKEKSDEDEDADELDEHKEA | 405 |
| CRTcolza | KEEEEKSK-----DAPAE TDGEDE-AEEDEGESDTESEKTE | 402 |
| CRTChlamydomonas_reinhardtii_a | EDEKAKK-----DAPPPVDAEAAEEDDEYEDKEEPSGM | 403 |
| CRTPhytophthora_infestans | QDEEKAK-----KEAAKAKAEKEDADEADEADEATA | 402 |
| CRTEhistolytica | EKQKEEE-----AKKEAEKQKEEETKKEIKKEENKE | 387 |
| CRT3_Humaine | EEMKKAR-----EEEEELLGKINRHEHYFNQFH | 379 |
| CRTtaureau | EEVKKAQ-----EEDEDDMLMGRFRGRENSEFKGFH | 379 |
| <hr style="border: 1px solid blue;"/> | | |
| Cter | | |
| CRTHumaine | AKDEL----- | 417 |
| CRTsouris | AKDEL----- | 416 |
| CRTXenopus | LKDEL----- | 416 |
| CRTsaumon | IKDEL----- | 421 |
| CRTthuitre | VKDEL----- | 414 |
| CRTHyalomma_anatolicum_excavat | KHEEL----- | 413 |
| CRTDermacentor_occidentalis_ti | KHEEL----- | 410 |
| CRTAmblyomma_americanum_tique | KHEEL----- | 410 |
| CRTHaemaphysalis_longicornis__ | KHEEL----- | 410 |
| CRTIxodes_ovatus_tique_ | DHEEL----- | 413 |
| CRTDrosophila_melanogaster | EHDEL----- | 406 |
| CRTOnchocerca_volvulus_nemato | ----- | ----- |
| CRTBrugia_malayi_nematode_ | THEEL----- | 421 |
| CRTMeloidogyne_incognita_nema | EHDEL----- | 415 |
| CRTTaenia_solium | SKDEL----- | 395 |
| CRTEchinococcus_granulosus_ten | PKDEL----- | 395 |
| CRTTcruzi | DKEDL----- | 401 |
| CRTTbrucei | DKSDL----- | 395 |
| CRTTrypanosoma | EEGDL----- | 397 |
| CRTMaïs | SLGRLP----- | 412 |
| CRTRicin | KYRKHYMD-----DYHDEL | 423 |
| CRTtabac | DKLES-----KDDEAHDEL | 416 |
| CRTsoja | SKTEAGED---TKEEGVHDEL | 420 |
| CRTPinus_taeda_pin_ | DKKEHELDSEHKEEDKKEHDEL | 427 |
| CRTcolza | AKSEVSEET--AEKDATAHDEL | 422 |
| CRTChlamydomonas_reinhardtii_a | GSIKIPKE----EEESGHDEL | 420 |
| CRTPhytophthora_infestans | QEAETKE----EEKTEKDEL | 419 |
| CRTEhistolytica | EL----- | 389 |
| CRT3_Humaine | RRNEL----- | 384 |
| CRTtaureau | RRNEF----- | 384 |
| <hr style="border: 1px solid blue;"/> | | |
| Cter | | |

ANNEXE 5 : ALIGNEMENT DE SEQUENCE DE CALRETICULINES

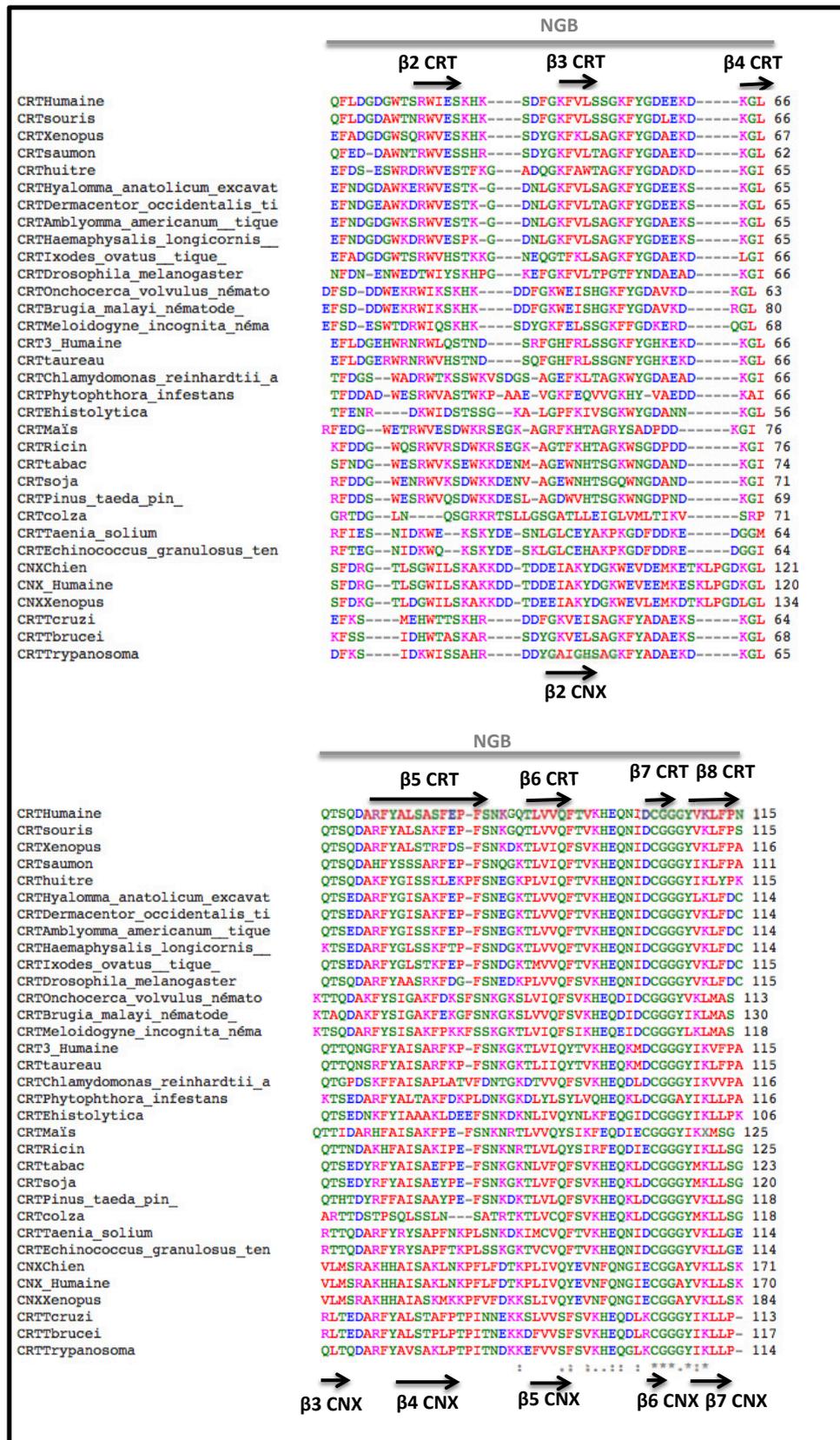
Alignement de séquence de calréticulines provenant de différentes espèces eucaryotes. Les domaines de la protéine, ainsi que les structures secondaires sont indiqués. Les résidus strictement conservés sont noté du symbole *.

| | | |
|--------------------------------|--|----|
| CRTHumaine | -----MLLSVPLLLGLLG----- | 13 |
| CRTsouris | -----MLLSVPLLLGLLG----- | 13 |
| CRTXenopus | -----MARVAFLLPLLAA----- | 14 |
| CRTsaumon | -----MLVSVLLMIALAS----- | 13 |
| CRThuitre | -----MKTVCVLLLLVG----- | 12 |
| CRTHyalomma_anatolicum_excavat | -----MRLLCILLPLVG----- | 12 |
| CRTDermacentor_occidentalis_ti | -----MRLLCILLPLVG----- | 12 |
| CRTAmblyomma_americanum_tique | -----MRLLCILLPLVG----- | 12 |
| CRTHaemaphysalis_longicornis__ | -----MRFLCILLPLVA----- | 12 |
| CRTIxodes_ovatus_tique_ | -----MRLVCVLLLLVG----- | 12 |
| CRTDrosophila_melanogaster | -----MMWCKTVIVLLATVG----- | 15 |
| CRTOnchocerca_volvulus_nemato | -----MQLSLLVGLVC----- | 11 |
| CRTBrugia_malayi_nematode_ | MREIDWKNEILGENSIEMQLYLLGLVC----- | 28 |
| CRTMeloidogyne_incognita_néma | -----MVSRCFCVSLLIACWP----- | 16 |
| CRT3_Humaine | -----MARALVQLWAICM----- | 13 |
| CRTtaureau | -----MAAARVPLWAICV----- | 13 |
| CRTChlamydomonas_reinhardtii_a | -----MKWGVVAVLATL----- | 12 |
| CRTPhytophthora_infestans | -----MLTKLVVALAAVA----- | 13 |
| CRTehistolytica | -----MFTLFL----- | 7 |
| CRTMaïs | -----MGTGRRGGGGGVFHRLLALSSL----- | 23 |
| CRTRicini | -----MGLLSLSELLITLTKLLPVFCLF----- | 23 |
| CRTtabac | -----MATQRRANP--SSLHLITVFSLL----- | 21 |
| CRTsoja | -----MAFRVRS--PDLRSLFLLSLL----- | 19 |
| CRTPinus_taeda_pin_ | -----MAGRRLS-----LYAVFLLLV----- | 17 |
| CRTcolza | -----MAKLTlass--PWFLSVSWRSLP----- | 21 |
| CRTTaenia_solium | -----MAVPALFIALLVFN----- | 15 |
| CRTechinococcus_granulosus_ten | -----MALLVFIALLFVVS----- | 15 |
| CNXChien | -----MEGKW-LLCMLLVGLTIVQAHEG-----HDD | 26 |
| CNX_Humaine | -----MEGKW-LLCMLLVGLTAIVEAHDG-----HDD | 26 |
| CNXXenopus | -----MDLKWFLVLTLLVHGLATINAHGGHHDDHDDHDDH | 37 |
| CRTTeruzi | -----MRAAIFFCALLGLAT----- | 15 |
| CRTTbrucei | -----MLMCMRPVAVACVFVALAT----- | 19 |
| CRTTrypanosoma | -----MRASLLLIALVGFAAV----- | 16 |

| | | |
|--------------------------------|--|----|
| CRTHumaine | -----LAVAEPAVYFKE | 25 |
| CRTsouris | -----LAAADPAIYFKE | 25 |
| CRTXenopus | -----LCIADPAVYFKE | 26 |
| CRTsaumon | -----AKPSVYFKE | 22 |
| CRThuitre | -----VVLSDPTEFFRE | 24 |
| CRTHyalomma_anatolicum_excavat | -----LISADPTVYFKE | 24 |
| CRTDermacentor_occidentalis_ti | -----LISADPTVYFKE | 24 |
| CRTAmblyomma_americanum_tique | -----LVAADPTVYFKE | 24 |
| CRTHaemaphysalis_longicornis__ | -----VASGDSNVYFKE | 24 |
| CRTIxodes_ovatus_tique_ | -----AVLADPKVYFKE | 24 |
| CRTDrosophila_melanogaster | -----FISAE--VYLKE | 25 |
| CRTOnchocerca_volvulus_nemato | -----FSAINAKIYFKE | 23 |
| CRTBrugia_malayi_nematode_ | -----FSAVSAKIYFKE | 40 |
| CRTMeloidogyne_incognita_néma | -----IFGVFGEVFFKE | 28 |
| CRT3_Humaine | -----LRVALATVYFQE | 25 |
| CRTtaureau | -----RRVALATVYFQE | 25 |
| CRTChlamydomonas_reinhardtii_a | -----VVAASAKDYFKE | 24 |
| CRTPhytophthora_infestans | -----ATVAQAEITFRE | 25 |
| CRTehistolytica | -----IALSSAKVYFHE | 19 |
| CRTMaïs | -----LLLASGEVVFEE | 35 |
| CRTRicini | -----VQFSVAEIIFEE | 35 |
| CRTtabac | -----VAVVSAEVFLEE | 33 |
| CRTsoja | -----S-IASANVFFEE | 30 |
| CRTPinus_taeda_pin_ | -----T-LVSAEVFFEE | 28 |
| CRTcolza | -----PLLSSRNALMMD | 33 |
| CRTTaenia_solium | -----TRAEVYFED | 24 |
| CRTechinococcus_granulosus_ten | -----SHAEVYFED | 24 |
| CNXChien | DMIDIEDDLLDDVIEEVEDSKK-PDTSAP-TSPKVTYKAPVPTGEVYFAD | 74 |
| CNX_Humaine | DVIDIEDDLLDDVIEEVEDSK--PDTPAPPSSPKVTYKAPVPTGEVYFAD | 73 |
| CNXXenopus | DHDHEDDGLDIDDELEDPEELKPEPSTPPPAPKVTYKAPVPTGEVYFSE | 87 |
| CRTTeruzi | -----LSAVHGTVYFHE | 27 |
| CRTTbrucei | -----VATVHGAIHFHE | 31 |
| CRTTrypanosoma | -----LTTVSATVLFHE | 28 |

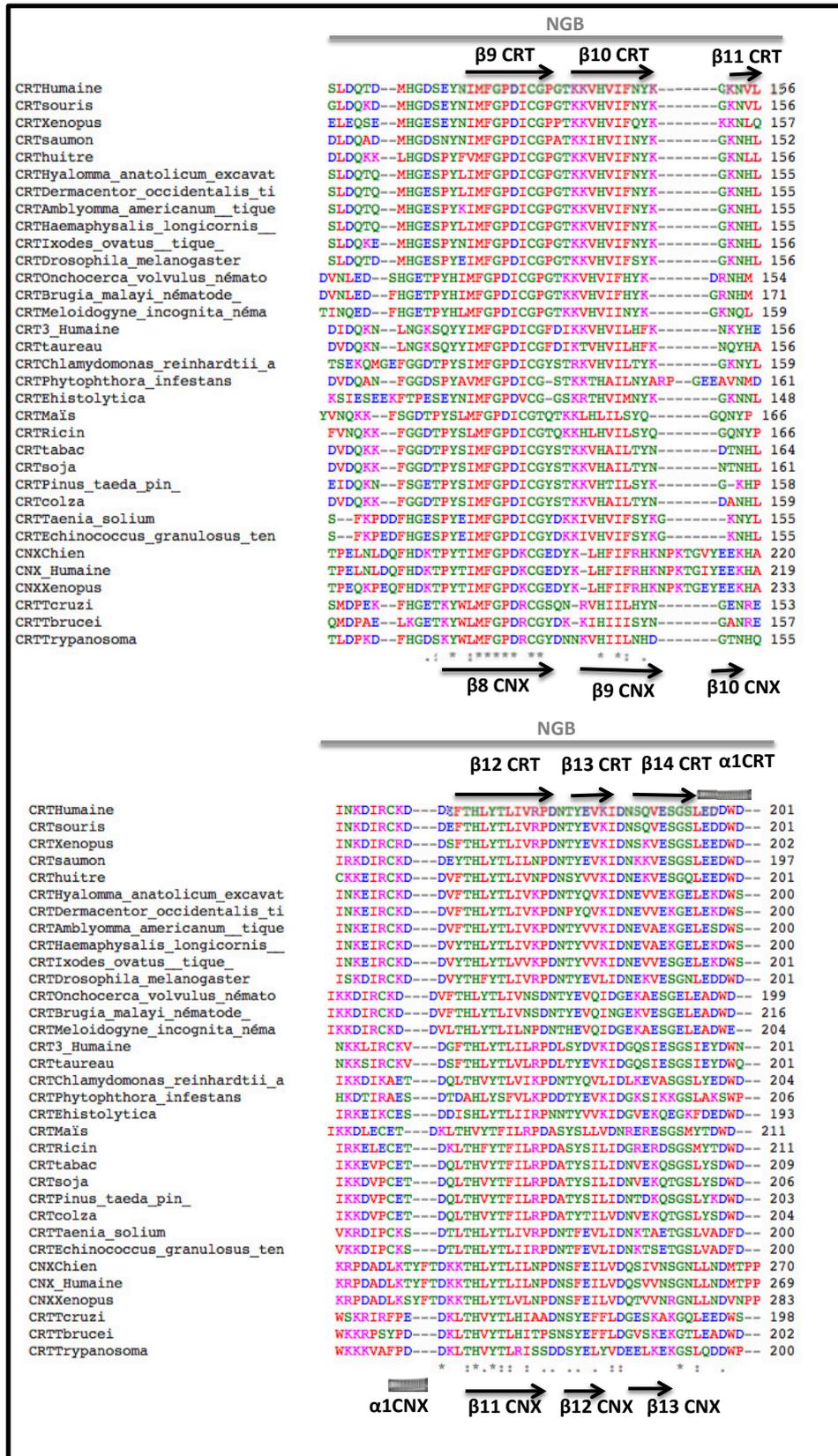
ANNEXE 6 : ALIGNEMENT DE SEQUENCE DE CALRETICULINES ET DE CALNEXINES

Alignement de séquence de calréticulines et de calnexines provenant de différentes espèces eucaryotes. Les domaines de la protéine, ainsi que les structures secondaires sont indiqués. Les résidus strictement conservés sont noté du symbole *



ANNEXE 6 : ALIGNEMENT DE SEQUENCE DE CALRETICULINES ET DE CALNEXINES

Alignement de séquence de calréticulines et de calnexines provenant de différentes espèces eucaryotes. Les domaines de la protéine, ainsi que les structures secondaires sont indiqués. Les résidus strictement conservés sont noté du symbole *.

**ANNEXE 6 : ALIGNEMENT DE SEQUENCE DE CALRETICULINES ET DE CALNEXINES**

Alignement de séquence de calréticulines et de calnexines provenant de différentes espèces eucaryotes. Les domaines de la protéine, ainsi que les structures secondaires sont indiqués. Les résidus strictement conservés sont noté du symbole *.

| | J | P | |
|--------------------------------|---|---------|---------|
| CRTHumaine | FLPPKKIKDPDASKPEDWDERAKIDDPDTSKPEDWDKP | ----- | 239 |
| CRTsouris | FLPPKKIKDPDAAKPEDWDERAKIDDPDTSKPEDWDKP | ----- | 239 |
| CRTXenopus | FLPAKKIKDPEAKKPDWDERPKIDDPDKKPEDWDKA | ----- | 240 |
| CRTsaumon | ILPPKKVKDPEAVKPDHWDERERMEDPDKKPEDWDRP | ----- | 235 |
| CRThuitre | FLPPKKIKDPEAKKPDWDEREKIDDPDKKPEDWDKP | ----- | 239 |
| CRTHyalomma_anatolicum_excavat | FLPPKKIKDPEAKKPEDWDDRAKIDDPDKKPEDWDKP | ----- | 238 |
| CRTDermacentor_occidentalis_ti | FLPPKKIKDPEAKKPEDWDDRAKIDDPDKKPEDWDKP | ----- | 238 |
| CRTAmblyomma_americanum_tique | FLPPKKIKDPEAKKPEDWDDRAKIDDPDKKPEDWDKP | ----- | 238 |
| CRTHaemaphysalis_longicornis__ | FLPPKIKDPAAKKPEDWDDRAKIDDPDKKPEDWDKP | ----- | 238 |
| CRTIxodes_ovatus_tique__ | LLPPKKIKDPAAKKPEDWDDRAKIDDPDKKPEDWDQP | ----- | 239 |
| CRTDrosophila_melanogaster | FLAPKKIKDPTATKPEDWDDRAITIPDDKKPEDWDKP | ----- | 239 |
| CRTOnchocerca_volvulus_nemato | FLPPKKIKDPAKKPEDWDEREFIDDEDKKPEDWDKP | ----- | 237 |
| CRTBrugia_malayi_nematode__ | FLPPKKIKDPAKKPEDWDEREYIDDEDKKPEDWDKP | ----- | 254 |
| CRTMeloidogyne_incognita_néma | LLPEKKIKDPAKKPEDWDETEYIDDPDKKPEDWDKP | ----- | 242 |
| CRT3_Humaine | LTSLKK-ETSPA-E-SKDWEQTK-----DNKAQDWEK | ----- | 230 |
| CRTtaureau | LTSLKKMEKASAE-AEGWDQAA-----KDKSQDWEK | ----- | 231 |
| CRTChlamydomonas_reinhardtii_a | MLPPKTIKDPKASKPEDWDEREETIADPEDKKPEGWDDIP | ----- | 243 |
| CRTPhytophthora_infestans | FQPEKQIKDPRQSKPKDQWDERQIADPADKKPEGWDDIP | ----- | 245 |
| CRTHistolytica | MLAPKEIDDPNVSKPADWVDEKEIDDPDKKPEGWDDIP | ----- | 232 |
| CRTMais | ILPPRKIKDVHAKRPKDWDDREYIEDPEVKKPEGYDSIP | ----- | 250 |
| CRTRicin | ILPPKKIRDTKAKKPADWDDREYIDDPNAVKKPEGYDKIP | ----- | 250 |
| CRTtabac | LLPPKTIKDPKASKPEDWDEKEFIDDPDKKPEGYDDIP | ----- | 248 |
| CRTsoja | LLPPKTIKDPKASKPEDWDDKEYIPDPDKKPEGYDDIP | ----- | 245 |
| CRTPinus_taeda_pin__ | LLPPKTIKDPNKKPEDWDDKEYIPDPDKKPEGYDDIP | ----- | 242 |
| CRTcolza | LLPPKKIKDPSAKKPEDWDEQEYIPDPDKKPDGYYDDIP | ----- | 243 |
| CRTAenia_solium | MIPPRTIDDPAKKPEDWVDEAIEIPDPDKKPDWDDQ | ----- | 237 |
| CRTechinococcus_granulosus_ten | MIPSKTIDDPAKKPEDWVVAEIEIPDPRKPDWDDQ | ----- | 237 |
| CNXChien | VNPSREIEDPEQKPEDWDERPKIPDPAVKPDDWDEDAPAKIPDEEATK | ----- | 320 |
| CNXHumaine | VNPSREIEDPEDRKPEDWDERPKIPDPAVKPDDWDEDAPAKIPDEEATK | ----- | 319 |
| CNXXenopus | VNPNREIEDPEDKKPEDWDERPKIPDPAVKPDDWDEDAPAKIPDENAVK | ----- | 333 |
| CRTTcruzi | LLLPRIVDETDKKPEDWVDEETMDPEDKKPEDWNEP | ----- | 237 |
| CRTTbrucei | FLPEKEIDDPEDKKPADWVDPITIDDPDKKPEDWSE | ----- | 241 |
| CRTTrypanosoma | IVQPKIISDATDKKPEDWDEPTMADPEDKKPEDWSE | ----- | 239 |
| | * : : * | | |
| | | β14 CNX | β15 CNX |
| | | P | |
| CRTHumaine | -----EHIPDPDAKKPEDWDEEMDGEWEP | ----- | 26 |
| CRTsouris | -----EHIPDPDAKKPEDWDEEMDGEWEP | ----- | 26 |
| CRTXenopus | -----EHIPDPDAVKPEDWDEEMDGEWEP | ----- | 26 |
| CRTsaumon | -----ENIADPDAKQPEDWDEEMDGEWEP | ----- | 26 |
| CRThuitre | -----EHIPDPDAKKPDWDEEMDGEWEP | ----- | 26 |
| CRTHyalomma_anatolicum_excavat | -----EYIPDPDATKPEDWDDMDGEWEP | ----- | 26 |
| CRTDermacentor_occidentalis_ti | -----EYIPDPDATKPEDWDDMDGEWEP | ----- | 26 |
| CRTAmblyomma_americanum_tique | -----EYIPDPDATKPEDWDDMDGEWEP | ----- | 26 |
| CRTHaemaphysalis_longicornis__ | -----EYIPDPDATKPEDWDDMDGEWEP | ----- | 26 |
| CRTIxodes_ovatus_tique__ | -----EYISDPDATKPEDWDEEMDGEWEP | ----- | 26 |
| CRTDrosophila_melanogaster | -----EHIPDPDATKPEDWDEEMDGEWEP | ----- | 26 |
| CRTOnchocerca_volvulus_nemato | -----EHIPDPDAKKPEDWDEEMDGEWEP | ----- | 262 |
| CRTBrugia_malayi_nematode__ | -----EHIPDPDAKKPEDWDEEMDGEWEP | ----- | 279 |
| CRTMeloidogyne_incognita_néma | -----ETIPDPDAKKPEDWDDMDGEWEP | ----- | 267 |
| CRT3_Humaine | -----HFLDASTSKQSDWNGDLGDWPAP | ----- | 25 |
| CRTtaureau | -----HFLDASASKPSDWKGLDGDWQAA | ----- | 25 |
| CRTChlamydomonas_reinhardtii_a | -----ATIADDAKKPEDWDEEDGTWEP | ----- | 26 |
| CRTPhytophthora_infestans | -----KTIIPDPDAKKPEDWDEEDGEWEP | ----- | 27 |
| CRTHistolytica | -----KTIIVDPNAKKPEEWNDEEDGEWEP | ----- | 25 |
| CRTMais | -----KQIPDPKDKKPDWDEEDGIWKPR | ----- | 275 |
| CRTRicin | -----KEIPDPNAKPEDWDEEDGIWKPR | ----- | 27 |
| CRTtabac | -----EEITDPDAKKPEDWDDQEDGEWTAP | ----- | 27 |
| CRTsoja | -----KELPDEAKKPEDWDEEDGEWTPP | ----- | 27 |
| CRTPinus_taeda_pin__ | -----KEIPDPDATKPEDWDEEDGEWTAP | ----- | 26 |
| CRTcolza | -----KEIPDPDAKKPEDWDEEDGEWTAP | ----- | 26 |
| CRTAenia_solium | -----PKTIIVDTNAKQPEDWNEETDGEWTAP | ----- | 26 |
| CRTechinococcus_granulosus_ten | -----PKTIIVDTNAKQPEDWNEETDGEWTAP | ----- | 26 |
| CNXChien | FDGWLDDPEYVDPDAEKPEDWDEMDGEWEPQIANPKCESAPGCGVW | ----- | 37 |
| CNXHumaine | FDGWLDDPEYVDPDAEKPEDWDEMDGEWEPQIANPKCESAPGCGVW | ----- | 36 |
| CNXXenopus | FDGWLDDPEYVDPDAEKPEDWDEMDGEWEPQVANPKCESAPGCGVW | ----- | 38 |
| CRTTcruzi | -----AMIPDADAKKPDWDAEDGFWEAP | ----- | 26 |
| CRTTbrucei | -----EKIVDPFAKKPEDWDAEDGFWEAP | ----- | 26 |
| CRTTrypanosoma | -----QTIPDAEAKKPEDWDAEDGFWEAP | ----- | 26 |
| | * : : * | | |

ANNEXE 6 : ALIGNEMENT DE SEQUENCE DE CALRETICULINES ET DE CALNEXINES

Alignement de séquence de calréticulines et de calnexines provenant de différentes espèces eucaryotes. Les domaines de la protéine, ainsi que les structures secondaires sont indiqués. Les résidus strictement conservés sont noté du symbole *.



ANNEXE 6 : ALIGNEMENT DE SEQUENCE DE CALRETICULINES ET DE CALNEXINES

Alignement de séquence de calréticulines et de calnexines provenant de différentes espèces eucaryotes. Les domaines de la protéine, ainsi que les structures secondaires sont indiqués. Les résidus strictement conservés sont noté du symbole *.

| | <u>Cter</u> | |
|--------------------------------|---|-----|
| CRTHumaine | FGVLGLDLWQVKS [*] GTIFDNFLITNDEAYAE [*] FGNETWGVTKAAEKQMK-- | 358 |
| CRTsouris | FAVLGLDLWQVKS [*] GTIFDNFLITNDEAYAE [*] FGNETWGVTKAAEKQMK-- | 358 |
| CRTXenopus | FGVLGLDLWQVKS [*] GTIFDNFLITDDEKYAE [*] YGNETWGVTKAAEKQMK-- | 359 |
| CRTsaumon | IGVIGLGLWQVKS [*] GTIFDNFLITDDATLAEVGNETWGVTKDPEKQMK-- | 354 |
| CRTthuitre | IGAVGFDLWQVKS [*] GTIFDNLLVTDDEKHAE [*] EVGNATWGVTKDPEKQMK-- | 358 |
| CRTHyalomma_anatolicum_excavat | ICKVGFDLWQVKS [*] GTIFDNLLITDDEEYARVHGEETWAAALRDAEKQMK-- | 357 |
| CRTDermacentor_occidentalis_ti | ICKVGFDLWQVKS [*] GTIFDNLLITDDEEYARVHGEETWAAALRDAEKQMK-- | 357 |
| CRTAmlyomma_americanum_tique | LCTVGFDLWQVKS [*] GTIFDNLLITDDEEYARVHGEETWAAALRDAEKQMK-- | 357 |
| CRTHaemaphysalis_longicornis__ | ICTVGFDLWQVKS [*] GTIFDNLLITDDEEYARVHGEETWAAALRDAEKQMK-- | 357 |
| CRTIxodes_ovatus_tique | ICTLGFDLWQVKS [*] GTIFDNLLVTDDEEYARVGEETWGVTKDGEKQMK-- | 358 |
| CRTDrosophila_melanogaster | ICTLGFDLWQVKS [*] GTIFDNVLTDDVELAAKAAAEVKN-TQAGEKQMK-- | 357 |
| CRTOnchocerca_volvulus_nemato | IGAIGFDLWQVKS [*] GTIFDDVIVTDSVEEAKKFG [*] EKTLKIREGEKQMK-- | 356 |
| CRTBrugia_malayi_nemato | IGAVGFDLWQVKS [*] GTIFDDVIVTDSVEEAKKFG [*] EKTLKIREGEKQMK-- | 373 |
| CRTMeloidogyne_incognita_nema | WGSVIGLWQVKS [*] GTIFDNIIVTDSIDEAKHAKETFEPLRDAEKQMK-- | 361 |
| CRT3_Humaine | IGAIGLELWQVRS [*] GTIFDNFLITDDEEYADNFGKATWGETKGP [*] EREMD-- | 344 |
| CRTtaureau | IGAVGLELWQVRS [*] GTIFDNFLITDDEEYADNFGKATWGETKGP [*] EREMD-- | 344 |
| CRTChlamydomonas_reinhardtii_a | LKRVGFELWQVKS [*] GSIFDNLLVTDDEEYARVHGEETWAAALRDAEKQMK-- | 362 |
| CRTPhytophthora_infestans | VGAIGFELWQVKS [*] GTIFDDLLVTDSEDDFNAHEEAVLAKT----EAMA-- | 361 |
| CRTehistolytica | FAYIGLWQVKS [*] GTIFDNLITDDEEYARVHGEETWAAALRDAEKQMK-- | 343 |
| CRTMaïs | LKYIGLEVWQVKS [*] GSVFDNLLICDDPEYARKVVEETWGANREAEKEAF-- | 369 |
| CRTRicin | LKYVGVVWQVKS [*] GSIFDNVLI [*] CDPEYAKVIVQEVF-ANREIEKEAF-- | 368 |
| CRTtabac | LKYVGVVWQVKS [*] GSIFDNVLI [*] CDPEYAKVIVQEVF-ANREIEKEAF-- | 367 |
| CRTsoja | LKYVGIELWQVKS [*] GSIFDNVLI [*] TDPEYAKVIVQEVF-ANREIEKEAF-- | 364 |
| CRTPinus_taeda_pin__ | LKYVGIELWQVKS [*] GSIFDNVLI [*] TDPEYAKVIVQEVF-ANREIEKEAF-- | 361 |
| CRTcolza | LKYVGIELWQVKS [*] GSIFDNVLI [*] CDPEYAKVIVQEVF-ANREIEKEAF-- | 362 |
| CRTTaenia_solium | FAYIGLGLWQVKS [*] GTIFDNFIVSDN--VSECQAHA [*] EYWKRF [*] AFE---- | 351 |
| CRTechinococcus_granulosus_ten | FAYIGLGLWQVKS [*] GTIFDNFIVSDN--VSECQAHA [*] EYWKRF [*] AFE---- | 351 |
| CNXChien | FSAIGLELWSM [*] TS [*] DI [*] FFDNFIVCGD [*] RRVVDWANDGWLKKAADGA [*] EPG | 469 |
| CNX_Humaine | FSAIGLELWSM [*] TS [*] DI [*] FFDNFIVCGD [*] RRVVDWANDGWLKKAADGA [*] EPG | 468 |
| CNXXenopus | FYAVGLELWSM [*] TS [*] DI [*] FFDNFIVT [*] SEK [*] RVADDWANDGWLKKAADGA [*] SAPS | 482 |
| CRTTeruzi | LTHVIGLWQVKS [*] GSIFK [*] DI [*] VI [*] GDD [*] LKEVLDLVEK [*] TYGGLKKA [*] EADAL-- | 357 |
| CRTBrucei | LTHVIGLWQVKS [*] GSIFK [*] DI [*] II [*] GDD [*] VKEVLDIVK [*] TYDGMKKA [*] EEDAL-- | 361 |
| CRTTrypanosoma | LEYVIGLWQVKS [*] GSIFN [*] II [*] I [*] GDD [*] VKEVLEIVK [*] TYGGLKKA [*] EADAL-- | 359 |
| | *:.:.*: :. :.:.: :. | |
| | <u>Cter CRT</u> | |
| CRTHumaine | -----DKQDEE | 364 |
| CRTsouris | -----DKQDEE | 364 |
| CRTXenopus | -----EQQDEE | 365 |
| CRTsaumon | -----ESQEEK | 360 |
| CRTthuitre | -----DEQDEE | 364 |
| CRTHyalomma_anatolicum_excavat | -----EKQEEE | 363 |
| CRTDermacentor_occidentalis_ti | -----DKQEEE | 363 |
| CRTAmlyomma_americanum_tique | -----EKQEEE | 363 |
| CRTHaemaphysalis_longicornis__ | -----DKQEEE | 363 |
| CRTIxodes_ovatus_tique | -----EQQDEK | 364 |
| CRTDrosophila_melanogaster | -----EAQDEV | 363 |
| CRTOnchocerca_volvulus_nemato | -----KK----- | 358 |
| CRTBrugia_malayi_nemato | -----EKQDE- | 378 |
| CRTMeloidogyne_incognita_nema | -----EAADEE | 367 |
| CRT3_Humaine | -----AIQAK- | 349 |
| CRTtaureau | -----AIQAK- | 349 |
| CRTChlamydomonas_reinhardtii_a | -----DKVKEE | 368 |
| CRTPhytophthora_infestans | -----AKKKQI | 367 |
| CRTehistolytica | -----AAEKKM | 349 |
| CRTMaïs | -----EEAEKE | 375 |
| CRTRicin | -----EEAEKI | 374 |
| CRTtabac | -----EEAEKK | 373 |
| CRTsoja | -----EEAEKK | 370 |
| CRTPinus_taeda_pin__ | -----DEAEKK | 367 |
| CRTcolza | -----DEIEKK | 368 |
| CRTTaenia_solium | ----- | |
| CRTechinococcus_granulosus_ten | ----- | |
| CNXChien | VVGQMI [*] EAAE [*] ERPWLWV [*] VYVLTVALPVFLVILFCCSGKKQSS [*] PVEY [*] KKTD | 519 |
| CNX_Humaine | VVGQMI [*] EAAE [*] ERPWLWV [*] VYVLTVALPVFLVILFCCSGKKQSS [*] PVEY [*] KKTD | 518 |
| CNXXenopus | VVGQMAAA [*] EERPWLWV [*] VYVLTVALPVFLVILFCCSGKKQPLDAE [*] HKKTD | 532 |
| CRTTeruzi | -----KVMEDM | 363 |
| CRTBrucei | -----AAFE-- | 365 |
| CRTTrypanosoma | -----EAFKKA | 365 |
| | ----- | |
| | Domaine cytoplasmique CNX | |

ANNEXE 6 : ALIGNEMENT DE SEQUENCE DE CALRETICULINES ET DE CALNEXINES

Alignement de séquence de calréticulines et de calnexines provenant de différentes espèces eucaryotes. Les domaines de la protéine, ainsi que les structures secondaires sont indiqués. Les résidus strictement conservés sont noté du symbole *.

| Cter CRT | |
|--------------------------------|---|
| CRTHumaine | QRLKEEEEKRRKKEE--EEAEDKEDDEDKDEDEEDEDKKEE----DEE 406 |
| CRTsouris | QRLKEEEEKRRKKEE--EEAEDKEDDDDRDEDEDEDEKKEE----DEE 406 |
| CRTXenopus | DRKKQEEEEKRRKDE--EPQEEEEEDDD--EEEEKEEKEE----EEE 405 |
| CRTsaumon | ERKKLEAEEEMARKEETKDEPGESEEEEEEELEHEEEEEDEEFGTGAQEE 410 |
| CRTthuitre | ERKKREEEKRRKKEE--EDNKKDDDDDDDEEGGDDDD----DEK 403 |
| CRTHyalomma_anatolicum_excavat | EEAKSKKEDD-AKDE-----DEFEEDEKDEKDEKDEEETTPAP- 403 |
| CRTDermacentor_occidentalis_ti | EEAKSKKEDD-AKDE-----DEFEEDEKDEKDEK--EETTPAP- 400 |
| CRTAmblyomma_americanum_tique | EDAKSKKEDD-AKDE-----DEFEEDEKDEKDEK--DEETTPAP- 400 |
| CRTHaemaphysalis_longicornis_ | EEAKNKKEDD-AKDE-----DEFDDDEDK--KEEE--EATTPAPS 399 |
| CRTIxodes_ovatus_tique | EDKEAEKDKK-EDDK-----DEFDEEDEDKKEE--EAPADKET 402 |
| CRTDrosophila_melanogaster | QRKKDEEEAKKASDK-----DDEDEDDDEEKDDE-----SKQ- 396 |
| CRTOnchocerca_volvulus_nemato | -----TKKQKKKEK-----NEKIKKE-----KMKRRKR-----ANR 384 |
| CRTBrugia_malayi_nematode_ | -----EKKKRRKEE-----EKKKKEEKEEEDKKEEKEE-----EDE 410 |
| CRTMeloidogyne_incognita_néma | ERKKFEEEEKRRKKEE--EESKKKDE-DKDGDEEEDKD-----EEE 404 |
| CRT3_Humaine | -----EEMKKAREE-----EEEELLSGKINR-----HEH 373 |
| CRTtaureau | -----EEVKAQEE-----DEDDMLMGRFRG-----REN 373 |
| CRTChlamydomonas_reinhardtii_a | -----EDEKKAADA-----PPPPVDAEAAEEDEEYEDKEEPPSGMGSIK 407 |
| CRTPhytophthora_infestans | -----QDEEKAKKE-----AAKAKEAEKEDADEADEAEDETAQEA 406 |
| CRTehistolytica | -----RDEIKAEAK-----QK-----EEAKKEAEKQKEE-----TKE 378 |
| CRTMaïs | -----RKAREDR-----EGQKAKDGGRRH-----IHRRHKKLQRS 407 |
| CRTRicin | -----RKAREEE-----EAQAREEGERRRRRGGYDRRHRDRYRDK 410 |
| CRTtabac | -----REEESK-----AAPADSDAE-----DDDADDAD--DADD 403 |
| CRTsoja | -----REEESK-----DDFVDSDAE-----EEDADEAGNDSDAES 403 |
| CRTPinus_taeda_pin_ | -----KEEEEKKE-----SDDEEDTDEKKEKSDDEDADELDDHEKAD 406 |
| CRTcolza | -----KEEEEK-----DAPAEFDGEDE--AEDEGESESDTESKTEA 403 |
| CRTaenia_solium | -----EEEEKKSFE-----EKDKESSTTDNVPDESEDEEVDLEEGDA 389 |
| CRTechinococcus_granulosus_ten | -----EEQEKKGFE-----EKESSTIESLPEESGGDEEVDLEEGDA 389 |
| CNXChien | APQPDPVKEEEEKKEE-----EKDKGDEEEEGEELKLEEKQKSDAEEDGGT 563 |
| CNX_Humaine | APQPDPVKEEEEKKEE-----EKDKGDEEEEGEELKLEEKQKSDAEEDGGT 562 |
| CNXXenopus | APQPDPVKEEEEDEE-----EFGKANKIEEEDAESEEPKQDEAEDQES 576 |
| CRTTcruzi | -----EKKRRKKEE-----EEEREKKAEE--EKKDEELEEKGDDGK 398 |
| CRTTbrucei | -----KKKEQDKKE-----ENKKEFDGEE--DKKKKE-----DK 392 |
| CRTTrypanosoma | -----EDAAAKAEE-----ANTATEVTIDNGDTKDAE-----DE 394 |
| Domaine cytoplasmique CNX | |
| Cter CRT | |
| CRTHumaine | EDVPGQAK-----DEL----- 417 |
| CRTsouris | E-SPGQAK-----DEL----- 416 |
| CRTXenopus | DEEETPLK-----DEL----- 416 |
| CRTsaumon | EEESDSIK-----DEL----- 421 |
| CRTthuitre | KDDEDKVK-----DEL----- 414 |
| CRTHyalomma_anatolicum_excavat | -DEDDHKH-----EEL----- 413 |
| CRTDermacentor_occidentalis_ti | -DDEDHKH-----EEL----- 410 |
| CRTAmblyomma_americanum_tique | -EDDDHKH-----EEL----- 410 |
| CRTHaemaphysalis_longicornis_ | EEDDDDHKH-----EEL----- 410 |
| CRTIxodes_ovatus_tique | EEEHQHDH-----EEL----- 413 |
| CRTDrosophila_melanogaster | -DKDQSEH-----DEL----- 406 |
| CRTOnchocerca_volvulus_nemato | KKKK----- 388 |
| CRTBrugia_malayi_nematode_ | KKKDDETH-----EEL----- 421 |
| CRTMeloidogyne_incognita_néma | EKKVVEEH-----DEL----- 415 |
| CRT3_Humaine | YFNQFHRR-----NEL----- 384 |
| CRTtaureau | SFKGFHRR-----NEF----- 384 |
| CRTChlamydomonas_reinhardtii_a | IPKEEEEES-----GHDEL----- 420 |
| CRTPhytophthora_infestans | ETKKEEKT-----EKDEL----- 419 |
| CRTehistolytica | EIKKEEN-----KEEL----- 389 |
| CRTMaïs | LGRLP----- 412 |
| CRTRicin | YRKHYMD-----DYHDEL----- 423 |
| CRTtabac | KLSE-----KDDEAHDEL----- 416 |
| CRTsoja | KTEAGED-----TKEGVHDEL----- 420 |
| CRTPinus_taeda_pin_ | KKEHELDSEHKEDDKHEDEL----- 427 |
| CRTcolza | KSEVSEET--AEKDATAHDEL----- 422 |
| CRTaenia_solium | SSKDEL----- 395 |
| CRTechinococcus_granulosus_ten | SPKDEL----- 395 |
| CNXChien | ASQEEDDRKPKAEDEEILNRSFPRNRPRE 593 |
| CNX_Humaine | VSQEEDDRKPKAEDEEILNRSFPRNRPRE 592 |
| CNXXenopus | QEEEEAEIEIKTKEDDIINRSFPRNRPRED 606 |
| CRTTcruzi | EDL----- 401 |
| CRTTbrucei | SDL----- 395 |
| CRTTrypanosoma | GDL----- 397 |
| Domaine cytoplasmique CNX | |

ANNEXE 6 : ALIGNEMENT DE SEQUENCE DE CALRETICULINES ET DE CALNEXINES

Alignement de séquence de calréticulines et de calnexines provenant de différentes espèces eucaryotes. Les domaines de la protéine, ainsi que les structures secondaires sont indiqués. Les résidus strictement conservés sont noté du symbole *.

RESUME

Au cours de cette thèse, nous avons étudié plus particulièrement deux calréticulines de parasite et leur interaction avec la protéine C1q humaine, dans un contexte de détournement du système immunitaire. En effet, une stratégie de mimétisme moléculaire avec des cellules apoptotiques a été proposée pour l'exposition de calréticuline à la surface de la forme infectieuse du parasite *Trypanosoma cruzi*, agent vecteur de la maladie de Chagas. Dans le cas du parasite *Entamoeba histolytica*, agent vecteur de l'amibiase, l'interaction de C1q avec la calréticuline exposée à la surface de l'amibe est utilisée pour mieux phagocyter les cellules de l'hôte.

La calréticuline est une protéine principalement localisée dans le réticulum endoplasmique où elle joue le rôle de protéine chaperonne en favorisant le repliement des protéines monoglycosylées. Une des fonctions extracellulaires de la calréticuline humaine est de favoriser l'élimination des cellules apoptotiques par les macrophages. Pour cela, la calréticuline est exposée à la surface des deux types cellulaires, à la surface desquelles elle est reconnue par C1q.

Nous avons résolu la structure de fragments des deux CRTs de parasite. Des interactions de type chaperonne et un aperçu de la flexibilité du domaine P ont été observés dans l'empilement cristallin et approfondis par analyse de diffusion des rayons X aux petits angles. Ces fragments générés pour l'analyse structurale nous ont permis en plus d'identifier une région clé dans l'interaction des calréticulines avec C1q. Du côté de C1q, deux mutations dans les régions globulaires de C1q (GRC1q), inhibent l'interaction avec la CRT et les IgM, suggérant un site partagé. Pour approfondir la caractérisation de l'interaction, des études du complexe par RMN ont débuté et nous avons développé une première forme recombinante monocaténaire de GRC1q, dont nous avons déterminé la structure. Nos recherches peuvent aider à développer des traitements contre ces parasites, aussi bien qu'à décrypter le rôle de la CRT des mammifères présente à la surface des macrophages et des cellules apoptotiques.

Mots clés: C1q, calréticuline, *Trypanosoma cruzi*, *Entamoeba histolytica*, cristallographie aux rayons X, immunité innée, subversion, GRC1q recombinantes.

SUMMARY

During this thesis, we were interested in two parasite calreticulin and their interaction with the human C1q protein in the context of the subversion of the immune system. Indeed, a molecular mimicry strategy with human apoptotic cells is suggested for the calreticulin exposure on the surface of the infectious form of the parasite *Trypanosoma cruzi*, which is responsible of Chagas' disease. In the case of the parasite *Entamoeba histolytica*, which is involved in amoebiasis, the interaction of C1q with the surface-exposed calreticulin is used to enhance phagocytosis of host cells.

Calreticulin is mainly localised in the endoplasmic reticulum, where it acts as a chaperone protein to favour the folding of monoglycosylated proteins. Moreover one of the extracellular functions of human calreticulin is to enhance the clearance of apoptotic cells by macrophages. This function is mediated through C1q interaction with calreticulin exposed on the surface of both cells.

We solved the structure of fragments of both parasite calreticulins. Chaperone-like interactions and an overview of the flexibility of the P domain were observed in the crystal packing and deepened using SAXS analyses. The fragments generated for X-ray crystallography studies allowed us to identify a key region of the interaction between C1q and the calreticulins. Two C1q mutations located in its globular regions (GRC1q) inhibit the interaction with calreticulin and IgM, suggesting a common binding area. To further characterise these interactions, we started NMR experiments and we produced the first single-chain recombinant form of GRC1q, which allowed solving its structure at high-resolution. Our investigations could provide tools to develop therapies against these parasites, and to decipher the role of mammal CRT on the surface of macrophages and apoptotic cells.

Keywords: C1q, calreticulin, *Trypanosoma cruzi*, *Entamoeba histolytica*, X-ray crystallography, innate immunity, pathogen subversion, recombinant GRC1q.