



HAL
open science

Exploration cristallographique de certains trésors cachés de la défense innée

Christine Gaboriaud

► **To cite this version:**

Christine Gaboriaud. Exploration cristallographique de certains trésors cachés de la défense innée. Biologie structurale [q-bio.BM]. Grenoble 1 UJF - Université Joseph Fourier, 2007. tel-01272548

HAL Id: tel-01272548

<https://hal.science/tel-01272548v1>

Submitted on 11 Feb 2016

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.

Mémoire en vue d'obtenir le
Diplôme d'Habilitation à Diriger les Recherches

Spécialité : Biologie
Université Joseph Fourier – Grenoble 1

Présenté par
Christine GABORIAUD

Laboratoire de Cristallographie
et Cristallogénèse des Protéines
Institut de Biologie Structurale Jean-Pierre Ebel

Date de soutenance : 06/09/07

Composition du jury :

Philippe Gasque	Rapporteur
Patrice Gouet	Rapporteur
Jean-Paul Mornon	Rapporteur
Christelle Breton	
Wim Bürmeister	
Gérard Arlaud	Collaborateur



*Ce qui ne sert apparemment à rien,
sert en réalité le plus.*

Maurice Bédart

*En hommage à Jules Bordet,
Prix nobel de Physiologie et de Médecine en 1919.*

Sommaire :

1-10 Curriculum vitae

11-14 **A. Transition après la thèse : protéases, analyses de séquence et modélisation.**

- A.1 Structures cristallographiques de protéases pancréatiques.
- A.2 Analyses de séquences et modélisation de protéines.

15- 32 **B. Introduction générale de la thématique « immunité innée ».**

- B.1. Mécanismes fondamentaux de protection de l'organisme.
- B.2. Présentation générale du système du complément.
- B.3 Protéines du complément : productions locales et autres fonctions
- B.4. Effets du système du complément relayés par les récepteurs cellulaires.
- B.5. Système du complément et pathologies.
- B.6. Informations structurales sur les protéines du complément.
- B.7 Les protagonistes de l'activation spécifique du complément.
- B.8. Les protéines de reconnaissance : C1q, MBL et ficolines.
- B.9. Structure modulaire commune des protéases associées : C1r, C1s et MASPs.
- B.10. Marqueurs moléculaires reconnus par C1q et les ficolines.
- B.11. Objectifs de l'étude structurale.

32-46 **C. Etudes structurales du complexe C1.**

- C. 1. Dissection et reconstitution du complexe C1 .
- C.2. Le domaine globulaire de reconnaissance de la protéine C1q.
- C.3. Domaine d'interaction de la protéase C1s.
- C.4. Domaine catalytique de la protéase C1r.
- C.5. Structure du domaine catalytique de C1s
- C.6. Conclusion sur la voie classique.

47-56 **D. Etudes structurales des complexes d'initiation de la voie lectine.**

- D.1. Domaines d'interaction des protéases MASPs.
- D.2. Domaine de reconnaissance des ficolines humaines.
- D3. Petite conclusion sur l'étude de la voie lectine.

57- 63 **E. Aperçu technique expérimental des projets.**

- E.1. Production-Purification
- E.2. Concentration :
- E.3. Cristallisation
- E.4. La résolution de la structure
- E.5. Diffusion de ligands dans des cristaux par trempage.

64-70 **F. Conclusions et perspectives.**

- F.1. Propriétés de reconnaissance de C1q et des ficolines.
- F.2. Assemblage et activation des complexes C1 et ficolines/MASPs.
- F.3. Une voie particulière de phagocytose des cellules apoptotiques ?
- F.4. Pathogènes et récepteurs du complément.

71-78 **Références**

Christine GABORIAUD

née le 1er Mai 1963 à Montmorency (95)

44 ans,

Française

Mariée, 3 enfants

Adresse professionnelle:

Institut de Biologie Structurale J.P. Ebel

Laboratoire de Cristallographie et Cristallogénèse des Protéines

UMR CEA-CNRS-UJF

41 rue Jules Horowitz

38027 Grenoble cedex 1

Tél: 04-38-78-95-99

Fax: 04-38-78-51-22

christine.gaboriaud@ibs.fr

Expérience professionnelle:

Situation actuelle (depuis octobre 1990)

Stage post-doctoral puis chercheur permanent au Commissariat à l'Energie Atomique, à l'Institut de Biologie Structurale J.-P. Ebel, au Laboratoire de Cristallographie et Cristallogénèse des Protéines dirigé par Juan Fontecilla-Camps à Grenoble.

Octobre 1989- septembre 90

Stage post-doctoral au laboratoire de Cristallographie (CNRS, Institut des Substances Naturelles, Gif sur Yvette), dans le groupe animé par Arnaud Ducruix. Contrat CEE. Sujet : Etude d'une protéase (papaïne) à site actif modifié.

Septembre 1987- septembre 1989

Doctorat (thèse soutenue le 15 septembre 1989, avec mention très honorable et les félicitations du jury) au Laboratoire de Minéralogie-Cristallographie (CNRS et Universités Paris 6 et 7, dirigé par Mr Petroff), sous la direction de Jean-Paul Mornon sur: " I. Etude cristallographique à 2 Å de résolution du complexe Trypsine de porc- Inhibiteur EETI 2. II. Développement d'une nouvelle méthode d'analyse de séquences de protéines: méthode HCA. Application à la modélisation par homologie."

Mai 1987 – Juillet 1987

Fin de scolarité à l'Ecole Normale Supérieure Ulm de Paris avec le stage de DEA: "Analyse statistique des textures des images de gels d'électrophorèse bidimensionnelle", effectué à l'ENS Ulm à Paris sous la direction de Philippe Tarroux

Encadrements:

Octobre 2006-

Co-encadrement de la thèse de Virginie Garlatti

Janvier -Juin 2005

Encadrement du stage de Master 2 BSN de Nicolas Belloy

Octobre 1999-Janvier 2004

Encadrement du DEA puis de la thèse de Lynn Gregory (Université J. Fourier, spécialité Cristallographie). « Etude structurale des protéases des complexes initiant la défense immunitaire par la voie du complément ».

2002-2003

Encadrement du stage post-doctoral de Jordi Juanhuix, impliqué dans la construction d'une ligne de lumière à l'ESRF, et désireux de se former à la biologie structurale, dans le cadre d'une collaboration à mi-temps.

2000- 2002

Encadrement du stage post-doctoral de Monika Budayova-Spano.

1997- 1998

Encadrement du stage de DEA (Cristallographie et RMN biologiques) de Cécile Olivier.

1992-1994

Participation à l'encadrement du doctorat de David Pignol

1993-1994

Encadrement du stage post-doctoral de Minxie Qing

1992-1993

Encadrement de la partie modélisation du DEA de Véronique Rossi (IBS/LEM). Cette collaboration s'est poursuivie pendant sa thèse sur l'étude cristallographique du domaine catalytique de C1s (sans succès), puis après sa thèse (enfin avec succès !).

1991-1992

Encadrement de la partie modélisation du DEA de Laurent David (avec Bertrand Toussaint et Paulette Vignais) .

Cours et enseignements:

Encadrante à l'école Franco-québécoise de Modélisation Macromoléculaire (6° et 7° cours de formation en biotechnologie). 16-29 août 1988 et 13-28 août 1989. Tignes (73)

Cours de 3h de formation sur la modélisation des protéines au Cours de Microbiologie Générale de l'Institut Pasteur (10 nov. 1995).

Cours au DEA de Biomathématiques à Paris VII.

Méthode HCA d'analyse de séquences d'acides aminés pour la modélisation macromoléculaire. Avril 1988

De la cristallographie à la modélisation des protéines. Avril 1989.

Jury de thèse :

Florence Teillet (Examineur) 22/09/06

Fabienne Dulin (Rapporteur) 19/12/06

Referee:

Cell Mol Life Sci., J. Mol. Biol. , Current Biology

Contrats :

Les contrats suivants ont été obtenus sur notre collaboration avec Gérard Arlaud (coordinateur):

Contrat BASF (1.4 MF pour les deux équipes)

Projet Structures et spécificité des modules CUB. Programme CNRS Physique et Chimie du Vivant.

Contrat ANR 2005 Microbiologie, Immunologie 38 000 € (pour ma partie)

Conférences et posters:

2nd International Conference on Crossroads between Innate and Adaptive Immunity, Crete, 17-22 Juin 2007.
'Innate immune sensing by human ficolins : Structural Insights'. Présentation orale courte + poster

XXIth International Complement Workshop, Pékin, Chine, 22_26 Octobre 2006.

Communications sélectionnées pour une présentation orale :

« The crystal structures of the binding domains of L- and H-ficolins provide insight into their recognition versatility », présentation C. Gaboriaud

- « The CUB1-EGF-CUB2 interaction domain of human MBL-associated serine protease-3 : X-ray structure and binding sites for MBL and ficolins », présentation N. Thielens
 « C1q binds phosphatidylserine and likely acts as an early bridging molecule in apoptotic cell recognition and clearance », présentation G. Arlaud.

Poster pour la Journée Technologie du 11 octobre 2006, à Minatec, Grenoble : « Les biosenseurs de l'immunité innée »

4th International Conference on Innate Immunity, Corfu, Grèce, 4-9 Juin 2006.

« The crystal structures of the binding domains of L- and H-ficolins reveal the versatile recognition mechanism of these innate immune proteins »

10th European Meeting on Complement in Human Disease , 9-13 Septembre 2005

« X-ray structures of the recognition domains of L- and H-ficolins reveal different carbohydrate binding sites »
 C. Gaboriaud, N. Belloy, L. Marti, M. Lacroix, M. Matsushita, T. Fujita, J.C. Fontecilla-Camps, G.J. Arlaud, N.M. Thielens

Séminaire invité : « The X-ray structure of L-ficolin and its GlcNAc binding site » Department of Medical Microbiology and Immunology, Université de Aarhus. Danemark. 5 novembre 2004.

3rd International conference on Innate Immunity, Aegean Conferences, Grèce, 10-15 Octobre 2004. Conférence intitulée: « X-ray structure of the trimeric fibrinogen-like domain of human L-ficolin and analysis of its carbohydrate binding site. »

20th International Complement Workshop, Honolulu, Hawaii, USA, 13-18 Juin 2004.

Conférence et posters intitulés: « The X-ray structure of human MBL-associated protein 19 (Map19) and its interaction site with mannan-binding lectin and L-ficolin ». Poster intitulé: « X-ray structure of the trimeric fibrinogen-like domain of human L-ficolin and analysis of its carbohydrate binding site. »

9th European Complement Workshop, Trieste, Italie, 6-9 septembre 2003. Conférence intitulée: « The crystal structure of the C1q globular domain unveils the basis of its versatile recognition properties ». Poster intitulé: « The X-ray structure of the CUB-EGF interaction domain of human protease C1s sheds light on the Ca²⁺ independent assembly of the C1 complex ».

Congrès International d'Immunologie, Rhodes, Grèce, 6-9 juin 2003. Poster intitulé : « Structural Analysis of the C1 complex : an overview ».

XIXth International Complement Workshop, Palerme, Italie, 22-26 septembre 2002.

Posters intitulés: " Crystal structure of the active catalytic domain of human C1r." et " The crystal structure of the zymogen catalytic domain of human C1r reveals that disruption of the C1r homodimer is required to trigger activation of the C1 complex."

Vth International Workshop on C1 and Collectins, Frankfurt, 26-28 octobre 2001. Poster intitulé: " The crystal structure of the zymogen catalytic domain of complement protease C1r reveals that a disruptive mechanical stress is required to trigger C1 activation."

Congrès de l'Association Française de Cristallographie, 3-6 juillet 2001, Orsay . Conférence intitulée: " Structure dimérique du domaine catalytique de la protéase C1r: un pas décisif vers la compréhension du mécanisme d'activation du complexe C1 du complément."

XIXth European Crystallographic Meeting, Nancy, France, 25-31 août 2000. Poster intitulé: "X-ray structure of the human C1s catalytic domain triggering the complement classical pathway: a trypsin-like domain modulated by a CCP module handle."

Société Novartis. Bâle, 22 août 2000. Conférence invitée: "X-ray structure of the human complement component C1s catalytic domain: a trypsin-like protease with a CCP module handle"

XVIIIth International Complement Workshop, Salt Lake City, Utah, USA; 23-27 juillet 2000. Conférence et poster intitulés: "X-ray structure of the human C1s catalytic domain: a protease with a CCP module handle."

VIIth DBMS/IBS Workshop « From the structure and function to the design of modular proteins » Autrans, France, 26-29 janvier 2000. Conférence intitulée: "Structural insights into the catalytic domain of complement protease C1s: an example of a rigid module-domain interactions."

Réunion de l'Association Française de Cristallographie (GTBIO99) Hourtin, France, 12-15 octobre 1999. Conférence intitulée: "Structure du domaine catalytique de C1s humain, initiateur de la cascade du complément."

Congrès de l'Association Française de Cristallographie, Orléans, France, 24-27 février 1998. Conférence intitulée: "Conservation à travers l'évolution d'une interaction rigide module-domaine dans les protéases modulaires."

IVth International Workshop on C1 and Collectins, Mayence, Allemagne, 3-5 octobre 1997. Posters intitulés: "Sequence conservation through evolution of the close interaction observed between the last CCP module and the serine protease domain of C1r and C1s." et "A three-dimensional model of the catalytic region of human C1r".

2nd International Conference on Molecular Structural Biology, Vienna, Autriche, 10-14 septembre 1997. Poster intitulé: "Conservation through evolution of the close interaction observed between the last CCP module and the serine protease domain of the C1r and C1s complement proteases"

48 heures de biologie structurale, Autrans, France, 23-24 mai 1996. Conférence intitulée: "Les protéases à sérine: un repliement bien connu mais des "finitions" accessibles seulement par une étude structurale approfondie."

2nd European Workshop on the crystallography of biological macromolecules, Come, Italie, 13-16 mai 1991. Poster intitulé: "X-ray structure of subunit III, a defective endopeptidase present in the bovine pancreatic procarboxypeptidase AS-6 complex".

15th International Crystallographic, Bordeaux, France, 19-28 juillet 1990. Poster intitulé: "X-ray Structure of the Porcine Trypsin-Inhibitor EETI 2 Complex at 2 Å resolution"

Réunion francophone de cristallographie des protéines, Paris, France, 21-23 mars 1990. Conférence intitulée: "Structure du complexe Trypsine de porc- inhibiteur EETI 2 à 2 Å de résolution"

Rencontres Biologie Moléculaire et Informatique (RBMI89), Paris, France, 8-9 juin 1989. Conférence invitée: "La méthode HCA: vers la prédiction de la similitude de repliements de certaines protéines à partir de l'analyse de leurs séquences."

Réunion conjointe de la British Biophysical Society et la Société Française de Biophysique sur la reconnaissance moléculaire, Carry-le Rouet, France, 12-14 avril 1989. Poster intitulé: "Crystallographical model of interaction between porcine Trypsin and EETI at 1.9 Å resolution."

Congrès franco-belge de Cristallographie des Macromolécules, Wégimont (Liège), Belgique, 19-21 septembre 1988. Posters intitulés: "Hydrophobic Cluster Analysis (HCA): une nouvelle méthode de comparaison et d'analyse des séquences de protéines." et "Cristallisation du complexe Trypsine de Porc / EETI et premier modèle structural à 3 Å de résolution."

3^o Congrès national de la Société Française de Chimie, Nice, France, 5-8 septembre 1988. Conférence invitée: "Modélisation moléculaire interactive 3D: réalités et perspectives dans le domaine de la recherche et de la formation."

Centre International de Recherches en Dermatologie, Sophia-Antipolis (06), 10 juin 1988, Conférence invitée: "Modèles d'interaction récepteurs-ligands".

Formation et diplômes:

Septembre 1989

Titre de docteur de l'Université Paris 7, spécialité Physique des matériaux, avec la mention très honorable et les félicitations du jury.

1983-1987, Ecole Normale Supérieure Ulm de Paris:

DEA de Biomathématiques (1986, Paris VII, mention TB)

Licence et maîtrise de Biochimie (1985, Paris VI, mention AB)

1981-1983

Classe préparatoire aux concours d'admission dans les grandes Ecoles, spécialité Biologie, au lycée Henri IV, Paris.

Juin 1981

Bac C, mention Bien.

Compétences diverses:

Informatique

Systèmes utilisés : VMS (Vax), Unix (Convex, DEC, SGI), Linux (PC).

Graphiques : stations Evans & Sutherlands, Silicon Graphics

Programmation en fortran 77, E&S,

Programmes utilisés en cristallographie et modélisation :

O, Turbo-Frodo, CCP4, PROTEIN, MERLOT, X-PLOR, CNS, AmoRe, Mosflm,

XDS, Denzo, Insight, Grasp, Hex

Administrateur système et réseau (Vax, Convex, SGI) de 1990 à 1992, installation de nombreux programmes pour la cristallographie pour le groupe LCCP.

Grands Instruments

Nombreuses expériences sur synchrotrons : LURE (Orsay, W32) et surtout ESRF (FIP, ID13, ID14, BM16, BM14).

Local contact occasionnel sur BM30

Publications:

REVUES :

1. Assembly of C1 and the MBL- and ficolin-MASP complexes: structural insights
C. Gaboriaud, F. Teillet, L.A. Gregory, N.M. Thielens, G.J. Arlaud (2007) *Immunobiol.*, 212, 279-288.
2. "Structure and activation of the C1 complex of complement: unravelling the puzzle". C. Gaboriaud, N. M. Thielens, L. A. Gregory, V. Rossi, J. C. Fontecilla-Camps and G. J. Arlaud. (2004) *Trends in Immunol.*, 25, 368-373.
3. " C1q and Tumor Necrosis Factor Superfamily:Modularity and Versatility " U. Kishore, C. Gaboriaud, P. Waters, A.K. Shrive, T.J. Greenhough, K.B. Reid, R.B. Sim, G.J.Arlaud (2004), *Trends in Immunol.*, 25, 551-561.
4. "Structural biology of C1. "Arlaud GJ, Gaboriaud C, Thielens NM, Rossi V. (2002) *Biochem. Soc. Trans.*. 100,1-6.
5. "Structural biology of the C1 complex of complement unveils the mechanisms of its activation and proteolytic activity. " G.J. Arlaud, C. Gaboriaud, N.M. Thielens, M .Budayova-Spano, V .Rossi, J.C. Fontecilla-Camps. (2002) *Mol Immunol.*, 39, 383-394.
6. "Structure, function and molecular genetics of human and murine C1r." G.J. Arlaud, C. Gaboriaud, G. Garnier, A. Circolo, N.M. Thielens, M. Budayova-Spano, J.C. Fontecilla-Camps, J.E. Volanakis. (2002) *Immunobiology*, 205,365-382
7. "Structural biology of C1: dissection of a complex molecular machinery." G. Arlaud, C. Gaboriaud, N. M. Thielens, V. Rossi, , B. Bersch & J.-F. Hernandez and J.C. Fontecilla-Camps. (2001) *Immunological Reviews*, 180, 136-145.
8. "Structural and functional studies on C1r and C1s: new insights into the mecanisms involved in C1 activity and assembly".G. J. Arlaud, V. Rossi, N. M. Thielens, C. Gaboriaud, B. Bersch & J.-F. Hernandez. (1998) *Immunobiol.*, 199, 303-316.
9. "Hydrophobic cluster analysis: procedures to derive structural and functional information from 2-D-representation of protein sequences." L. Lemesle-Varloot, B. Henrissat, C. Gaboriaud, V. Bissery, A. Morgat and J.P. Mornon (1990) *Biochimie* 72, 555-74.

RESULTATS ORIGINAUX :

10. "Structural insights into the innate immune recognition specificities of L- and H-ficolins". V. Garlatti, N. Belloy, L. Martin, M. Lacroix, M. Matsushita, Y. Endo, T. Fujita, J.C. Fontecilla-Camps, G.J. Arlaud, N.M. Thielens, C. Gaboriaud (2007) *EMBO J.*, 26, 623-633.

11. “ Functional role of the linker between the complement control protein modules of complement protease C1s. ” I. Bally, V. Rossi, N.M. Thielens, C. Gaboriaud, G.J. Arlaud (2005) *J. Immunol.* 175, 4536-4542.

12. “The major MBL-I and II forms of human mannan-binding lectin : chemical characterization, carbohydrate binding properties and interaction with MBL-associated serine proteases. ” F. Teillet, B. Dublet, J.-P. Andrieu, C. Gaboriaud, G.J. Arlaud, N. Thielens. (2005) *J. Immunol.*, 174, 2870-2877.

13. “The X-ray structure of human Mannan-binding lectin- Associated protein 19 (MAp19) and its interaction site with Mannan-binding Lectin and L-ficolin. ” L.A. Gregory, N.M. Thielens, M. Matsushita, R. Sorensen, G.J. Arlaud, , J.C. Fontecilla-Camps & C. Gaboriaud (2004), *J. Biol. Chem.*, 279, 29391-29397.

14. “The crystal structure of the globular head of complement protein C1q provides a basis for its versatile recognition properties. ” C. Gaboriaud, J. Juanhuix, A. Gruez, M. Lacroix, C. Darnault, D. Pignol, D. Verger, J.C. Fontecilla-Camps, G.J. Arlaud (2003) *J. Biol. Chem.* 278, 46974-46982.

15. “X-ray structure of the Ca²⁺-binding interaction domain of C1s. Insights into the assembly of the C1 complex of complement. ” L.A. Gregory, N.M. Thielens, G.J. Arlaud, J.C. Fontecilla-Camps, C. Gaboriaud (2003) *J. Biol. Chem.*, 278, 32157-32164.

16. “Monomeric structures of the zymogen and active catalytic domain of complement protease C1r: further insights into the C1 activation mechanism.” M. Budayova-Spano, W. Grabarse N.M. Thielens, H. Hillen, M. Lacroix, M. Schmidt, J.C. Fontecilla-Camps, G.J. Arlaud, C. Gaboriaud. (2002) *Structure*, 10, 1509-1519.

17. “ 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. ” M. Budayova-Spano, M. Lacroix, N. M. Thielens, G. J. Arlaud, J. C. Fontecilla-Camps and C. Gaboriaud (2002) *EMBO J.*, 21, 231-239

18. “ Crystal structure of the catalytic domain of human complement C1s: a serine protease with a handle .”. C. Gaboriaud, V. Rossi, I. Bally, G. Arlaud, and J.C. Fontecilla-Camps. (2000) *EMBO J.*, 19, 1755-1765.

19. “Evolutionary conserved rigid module-domain interactions can be detected at the sequence level: the examples of complement and blood coagulation proteases ”. C. Gaboriaud, V. Rossi, J.-C. Fontecilla-Camps& G. Arlaud. (1998) *J. Mol. Biol.* 282, 549-470.

20. "NMR structures of the C-terminal end of human complement serine protease C1s." P. Gans, V. Rossi, C. Gaboriaud, I. Bally, J.-F. Hernandez, M.J. Blackledge & G.J. Arlaud (1998) *Cell Mol. Life Sci.*, 54, 171-178.
21. "Structure and assembly of the catalytic region of human complement protease C1r: a three-dimensional model based on chemical cross-linking and homology modeling". M. Lacroix, V. Rossi, C. Gaboriaud, S. Chevallier, M. Jaquinod, N.M. Thielens, J. Gagnon & G. Arlaud. (1997) *Biochemistry* 36, 6270-6282.
22. "Crystal structure of human trypsin 1: unexpected phosphorylation of Tyr 151". C. Gaboriaud, L. Serre, O. Guy-Crotte, E. Forest and J.-C. Fontecilla-Camps. (1996) *J. Mol. Biol.* 259, 995-1010.
23. "How to escape from the model bias with a high resolution native data set: structure determination of the PcpA-S6 subunit III." D. Pignol, C. Gaboriaud, J.C. Fontecilla-Camps, V.S.Lamzin and K.S.Wilson. (1996) *Acta Cryst. D* 52, 345-355.
24. "Structure of the catalytic region of human complement protease C1s: study by chemical cross-linking and three-dimensional homology modeling." V. Rossi, C. Gaboriaud, M. Lacroix, J. Ulrich, J.-C. Fontecilla-Camps, J. Gagnon and G. J. Arlaud (1995). *Biochemistry* 34, 7311-7321.
25. "Crystal structure of bovine procarboxypeptidase A-S6 subunit III, a highly structured truncated zymogen E." D. Pignol, C. Gaboriaud, T. Michon, B. Kerfelec, C. Chapus and J.-C. Fontecilla-Camps (1994) *EMBO J.* 13, 1763-1771.
26. "Mansek and Sunhca - 2 interactive programs for the Hydrophobic Cluster Analysis of Protein Sequences" L. Lemesle-Varloot, C. Gaboriaud, A. Morgat, G. Pantel, J.-P. Mornon, S. Lavaitte, F. Lestang & B. Henrissat. (1993) *Comp. Appl. in the Biosciences*, 9, 37-44
27. "Protein design on computers. Five new proteins: Shpilka, Grendel, Fingerclasp, Leather and Aida." Sander C, Vriend G, Bazan F, Horovitz A, Nakamura H, Ribas L, Finkelstein AV, Lockhart A, Merkl R, Perry LJ, et al.(1992) *Proteins: Struct., Funct., Genet.* 12, 105-110.
28. "Hydrophobic cluster analysis reveals duplication in the external structure of human α -interferon receptor and homology with γ -interferon receptor external domain". C. Gaboriaud, G. Uzé, G. Lutfalla and K. Mogensen. (1990) *FEBS Lett.* (1990) 269, 1-3.

29. "Use of Restrained Molecular Dynamics in Water to Determine 3D Protein Structure: Prediction of the 3D structure of Trypsin Inhibitor EETI-II." L. Chiche L., C. Gaboriaud, A. Heitz, J.P. Mornon, B. Castro and P.A. Kollman (1989). *Proteins: Struct., Funct., Genet.* 6, 405-417.
30. "Crystallisation and preliminary X-ray study of porcine Trypsin, free and complexed with *Ecballium Elaterium* Trypsin Inhibitor." C. Gaboriaud, M.C. Vaney, B. Bachet, D. Le-Nguyen, B. Castro and J.P. Mornon. (1989) *J. Mol. Biol.*, 210, 883-884.
31. "Hydrophobic Cluster Analysis (HCA) of the hormone-binding domain of receptor proteins." J.P. Mornon, V. Bissery, C. Gaboriaud, A. Thomas, T. Ojasoo and J.P. Raynaud. (1989) *J. Ster. Biochem.* 34, 355-361.
32. "Hydrophobic Cluster Analysis: A new way to analyse and compare amino-acid sequences." C. Gaboriaud, V. Bissery, T. Benchetrit and J.P. Mornon. (1987) *FEBS Lett.*, 224, 149-155.
33. "HERMeS: A second generation approach to the automatic analysis of two-dimensional electrophoresis gels: Part I: Data acquisition." P. Vincens, N. Paris, J.L. Pujol, C. Gaboriaud, T. Rabilloud, J.L. Pennetier, P. Matherat, P. Tarrowx. (1986) *Electrophoresis*, 7, 347-356.

OUVRAGES :

34. « C1q: a versatile pattern recognition molecule and sensor of altered self » (2007) C. Gaboriaud, H. Païdassi, P. Frachet, P. Tacnet-Delorme, N.M. Thielens, G. J. Arlaud. Sous presse dans « Collagen-related lectins in innate immunity », D. Kilpatrick Ed.
35. « The classical pathway C1 complex » G.J. Arlaud, C. Gaboriaud, N. Thielens, L.A. Gregory, J. Juanhuix, V. Rossi, J.C. Fontecilla-Camps. (2004) pp 63-89 in « Structural Biology of the Complement System », D. Morikis & J. Lambris, Eds, CRC Press..
36. « Complement component C1r » G.J. Arlaud, N. Thielens, C. Gaboriaud, (2004) pp1617-1620 in « Handbook of Proteolytic Enzymes, 2nd edition » , Elsevier, ISBN 0-12-079610-4
37. « Complement component C1s » G.J. Arlaud, V. Rossi, N. Thielens, C. Gaboriaud, (2004) pp1620-1623 in « Handbook of Proteolytic Enzymes, 2nd edition » , Elsevier, ISBN 0-12-079610-4
38. "An analysis of the steroid binding domain of receptors and of ligand structure and binding affinity." J.P. Raynaud, V. Bissery, C. Gaboriaud, T. Ojasoo, G. Teutsh, J.P. Mornon. (1989) in "The Steroid/Thyroid Hormone Receptor Family and Gene Regulation" (J. Carlstedt-Duke, H. Eriksson, J.A. Gustafsson, Eds) Life Sciences, vol. 4, pp.337-366. Birkhauser Verlag, Basel.

A. Transition après la thèse : protéases, analyses de séquence et modélisation par homologie.

La biologie structurale, en particulier la cristallographie des protéines, est au centre de mon activité scientifique depuis ma thèse. Après une formation en biochimie et biomathématiques à l'École Normale Supérieure de Paris, la thèse a été pour moi l'occasion de découvrir plus en profondeur l'univers des protéines, et plus particulièrement les relations entre leurs séquences et leurs structures, ou leurs structures et leurs fonctions. Par la suite, j'ai orienté mon travail en privilégiant l'utilisation des informations expérimentales obtenues par cristallographie et diffraction des rayons X, les relations séquence-structure-fonction étant toujours considérées pour l'étape de valorisation des structures obtenues.

A l'heure actuelle, mon travail de recherche est centré sur la thématique de l'immunité innée, plus particulièrement sur l'étude des mécanismes de reconnaissance des éléments dangereux et sur le déclenchement de mécanismes effecteurs de défense induits par cette reconnaissance initiale. Ce travail est le fruit d'une collaboration étroite avec le groupe de Gérard Arlaud, collaboration interne à l'Institut de Biologie Structurale (IBS). Les différents conseils scientifiques de l'IBS ont souligné l'importance de telles collaborations internes pour la dynamique et la lisibilité de l'activité de l'Institut de Biologie Structurale.

Ce document commence par une description très sommaire des travaux réalisés depuis la thèse, soutenue en 1989. Ceci permettra de mieux apprécier la continuité de mon parcours de recherche et illustrera la diversité des questions abordées et des méthodes utilisées. La thématique actuelle et les projets associés feront l'objet d'une présentation plus détaillée.

A.1 Structures cristallographiques de protéases pancréatiques.

Ayant résolu au cours de ma thèse la structure cristallographique d'un complexe trypsine-inhibiteur, j'ai commencé à travailler naturellement dans le groupe de Juan Fontecilla-Camps sur la thématique « structures de protéases pancréatiques », en collaboration avec des équipes de biochimistes à Marseille.

A.1.1. Structure de la trypsine humaine de type 1 : modification de la tyrosine 151.

(avec O. Guy et C. Figarella, groupe de recherche sur les glandes exocrines, Faculté de Médecine de Marseille ; et E. Forest, laboratoire de spectrométrie de masse de l'IBS)

L'étude structurale de cette enzyme était motivée par l'observation de propriétés biochimiques particulières et inexplicables. La structure, résolue et affinée à 2.2 Å de résolution, a donné deux résultats inattendus : a) une tyrosine modifiée à 10 Å du site actif ; b) un empilement cristallin suggérant une association sous forme de tétramères.

Une forme tétraédrique de la densité électronique d'une tyrosine (Figure 1) nous a fait envisager soit une modification post-traductionnelle, soit un site secondaire de fixation du DFP, utilisé de façon massive pour limiter l'autoprotéolyse. L'analyse d'un échantillon de trypsine humaine native (non traitée au DFP) par spectrométrie de masse a permis de montrer que cette modification n'était pas un artefact lié à l'utilisation massive de DFP, et la masse obtenue était compatible avec l'hypothèse d'une phosphorylation. Une étude très récente vient de montrer qu'il s'agit en fait d'une sulfatation de la tyrosine, dont le rôle physiologique n'est pas encore très clair, si ce n'est qu'il accélère la vitesse d'autoactivation pour l'espèce cationique (Sahin-Toth et al., 2006). Cette modification n'est retrouvée que chez les primates. Elle est absente dans la trypsine de bœuf, ainsi que dans les formes recombinantes de trypsine humaine. Cette étude

structurale nous a donc permis de réaliser la première observation de ce type dans cette classe d'enzyme. De plus, une analyse comparative minutieuse avait permis de corrélérer

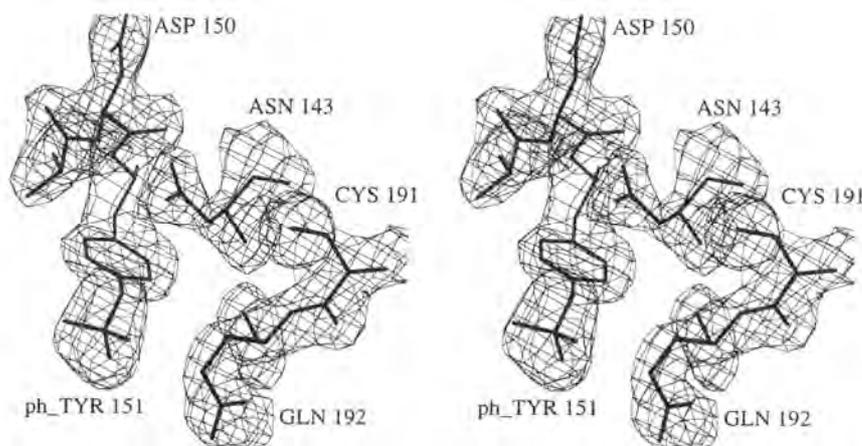


Figure A1 :

Vue stéréoscopique du modèle de la trypsine humaine au niveau de la tyrosine modifiée. La carte de densité électronique 2Fo-Fc est contournée à 1 σ .

les particularités biochimiques, comme la résistance à plusieurs inhibiteurs spécifiques des trypsines, à des particularités structurales, dont cette modification inattendue.

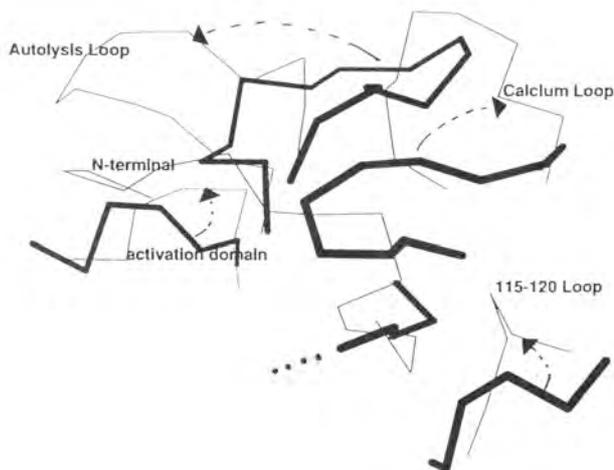
Publi 22 : C. Gaboriaud, L. Serre, O. Guy-Crotte, E. Forest and J.-C. Fontecilla-Camps (1996) Crystal structure of human trypsin 1: unexpected phosphorylation of Tyr 151. *J. Mol. Biol.*, **259**, 995-1010.

A.1.2 Sous-unité III du complexe procarboxypeptidase-A S6: la structure ordonnée inattendue d'un zymogène tronqué en N-terminal.

(avec D. Pignol, et en collaboration avec le groupe de C. Chapus, Centre de Biochimie et de Biologie Moléculaire du CNRS, Marseille)

Cette sous-unité du complexe ternaire Procarboxypeptidase –A S6, sécrétée par le pancréas de boeuf, est une endopeptidase qui apparaît inactive et tronquée en N-terminal après purification. La structure a permis de révéler les changements structuraux majeurs liés à l'absence en N-terminal du dipeptide hydrophobe extrêmement conservé dans les protéases à sérine : bien que 80% du modèle présente une architecture classique de type trypsine, l'orientation des boucles dites d'autolyse et de fixation du calcium sont dans des conformations très différentes de celles déjà observées dans les protéases à sérine. Le site de fixation du substrat est bloqué en conformation inactive. Le passage à la forme active impliquerait de larges mouvements concertés, peu vraisemblables vu le caractère très ordonné de la structure. Cette forme atypique de zymogène révèle le rôle déterminant du dipeptide hydrophobe dans le processus d'activation.

Sur le plan méthodologique, des problèmes ont été rencontrés pour résoudre cette structure par remplacement moléculaire, à cause du fort biais initial du modèle au niveau des boucles de surface dont la conformation a été fortement modifiée dans cette protéase atypique. Ceci m'a conduit à implanter au laboratoire la méthode ARP développée par V. Lamzin et K. Wilson, qui nous a permis de finir de résoudre la structure de la procarboxypeptidase-A S6.



En l'absence du dipeptide hydrophobe en N-terminal, une conformation atypique des boucles de surface conduit à l'inactivité de la molécule.

Conformation observée pour les boucles de surface (en gras) et mouvements nécessaires pour rétablir une forme active (traits fins)

Publi 25 : D. Pignol, C. Gaboriaud, T. Michon, B. Kerfelec, C. Chapus and J.-C. Fontecilla-Camps. (1994) Crystal structure of bovine procarboxypeptidase A-S6 subunit III, a highly structured truncated zymogen E. *EMBO J.*, **13**, 1763-71.

Publi 23 : D. Pignol, C. Gaboriaud, J.C. Fontecilla-Camps, V.S.Lamzin and K.S.Wilson. (1996) How to escape from the model bias with a high resolution native data set: structure determination of the PcpA-S6 subunit III *Acta Crystallographica D* **52**, 345-355.

A.2 Analyses de séquences et modélisation de protéines.

Le deuxième volet de ma thèse concernait la modélisation des protéines, avec le développement de la méthode d'analyse de séquences HCA (Hydrophobic Cluster Analysis, publi n° 32). Cette méthode repose sur des principes très différents des méthodes classiques d'alignement de séquences (type Blast). Elle repose sur une représentation graphique originale de la séquence des protéines, qui met en valeur le regroupement des acides aminés hydrophobes associés aux faces internes des structures secondaires régulières. Cette méthode est adaptée à l'analyse de similarités structurales (par conservation du coeur hydrophobe) de protéines, même à faible similarité de séquence. L'amélioration et l'exploitation de cette méthode se sont poursuivies et ont donné lieu à plusieurs centaines de publications. Je n'ai que ponctuellement poursuivi cette thématique en marge de mon activité principale de cristallographe des protéines. Cependant, mon expérience en alignement de séquences, alignement de structures et en modélisation des protéines a toujours été un atout supplémentaire, et parfois décisif, pour mener à bien mes projets.

A.2.1 Modélisation par homologie de protéines.

Avant de réussir à obtenir la première structure cristallographique sur le projet « immunité », nous avons utilisé la modélisation par homologie (à partir des structures de protéines de la même famille) pour construire les premiers modèles des protéases C1r et C1s, en tenant compte d'informations expérimentales diverses comme l'analyse par réticulation chimique, les images de microscopie électronique, etc... Cette technique offre un très bon support pour initier la discussion entre « biochimistes » et « structuralistes » sur l'étude des relations structure-fonction. J'ai formé deux étudiants de DEA à la modélisation par homologie : Laurent David pour la modélisation de IHF et plus particulièrement Véronique Rossi, dans le groupe de Gérard Arlaud, pour la modélisation des fragments catalytiques γ B de C1r et C1s.

- Publi 24 : V. Rossi, C. Gaboriaud, M. Lacroix, J. Ulrich, J.-C. Fontecilla-Camps, J. Gagnon and G. J. Arlaud. (1995) Structure of the catalytic region of human complement protease C1s: study by chemical cross-linking and three-dimensional homology modeling. *Biochemistry* 34, 7311-21.
- Publi 21 : M. Lacroix, V. Rossi, C. Gaboriaud, S. Chevallier, M. Jaquinod, N.M. Thielens, J. Gagnon & G. Arlaud. (1997) Structure and assembly of the catalytic region of human complement protease C1r: a three-dimensional model based on chemical cross-linking and homology modeling *Biochemistry* 36, 6270-6282.

A.2.2 Exploitation de la méthode d'analyses de séquences HCA.

Pour ce qui est de l'utilisation de la méthode HCA, je n'ai pu que conseiller ponctuellement d'autres chercheurs sur des problèmes de reconnaissance d'homologie structurale ou d'alignement de séquences.

- Publi 28 : C. Gaboriaud, G. Uzé, G. Lutfalla & K. Mogensen (1990) Hydrophobic cluster analysis reveals duplication in the external structure of human α -interferon receptor and homology with γ -interferon receptor external domain. *FEBS Lett.* 269, 1-3.

J'ai aussi utilisé la méthode HCA lors de la modélisation des modules CCP des fragments γ B de C1s et C1r (voir ci-dessus).

A.2.3 Prédiction de la conservation d'un assemblage rigide d'un module sur un domaine par une stratégie originale d'analyse comparative de séquences.

J'ai développé une méthode de comparaison d'alignements de séquences multiples qui m'a permis de prédire l'existence d'une interface conservée entre le domaine SP et le module CCP des protéases du système du complément. Cette méthode permettait même de prédire les acides aminés conservés à cette interface. De façon remarquable, ces résidus ont été confirmés ensuite, à une seule exception près, par l'analyse cristallographique des structures de C1s, C1r et MASP2.

- Publi 19 : C. Gaboriaud, V. Rossi, J.-C. Fontecilla-Camps & G. Arlaud. (1998) Evolutionary conserved rigid module-domain interactions can be detected at the sequence level: the examples of complement and blood coagulation proteases. *J. Mol. Biol.* 282, 549-470.

A.2.4 Analyses de séquences, comparaisons de structure et cristallographie des protéines.

L'étape d'analyse et de valorisation d'une structure implique souvent un travail de modélisation pour intégrer, dans l'espace 3D, l'ensemble des informations expérimentales disponibles et proposer des hypothèses fonctionnelles. Un bon savoir-faire dans la comparaison des séquences et des structures est également un atout. Enfin, comme cela sera évoqué brièvement au chapitre E, une bonne maîtrise et utilisation des alignements de séquences peut être cruciale pour résoudre une structure par remplacement moléculaire (par exemple pour les domaines globulaires de C1q).

B. Introduction générale de la thématique « immunité innée ».

Les études structurales développées au cours des 10 dernières années s'inscrivent dans le cadre de l'analyse des relations structure-fonction de molécules du système immunitaire, et plus particulièrement, du système du complément. Les systèmes de défense utilisés chez l'homme ont atteint au cours de l'évolution un fort degré de complexité et de redondance. Ces procédés de défense mettent en œuvre des cellules, dont le développement est plus ou moins spécialisé mais aussi des molécules présentes dans le milieu extracellulaire. On peut distinguer l'immunité innée, disponible immédiatement à la naissance, et dont les actions sont immédiates. Elle met en jeu des processus acquis assez tôt dans l'évolution, avant ceux de l'immunité dite 'adaptative'. Ces derniers sont mis en œuvre par des cellules lymphocytaires spécialisées (défense cellulaire), dont la sélection et l'apprentissage sont réalisés plus progressivement, pour obtenir à terme une action plus spécifique et efficace, en synergie avec les autres mécanismes de défense.

Une meilleure compréhension de l'ensemble des mécanismes impliqués devrait nous aider à définir des stratégies d'intervention en cas de déficience, de perturbation ou dérèglement d'une partie du système, voire en cas de subversion par des pathogènes ou des cellules cancéreuses.

B.1. Mécanismes fondamentaux de protection de l'organisme.

L'organisme peut être confronté à des agressions extérieures (agents infectieux) et des dérèglements internes (par exemple cancer, maladies autoimmunes).

Certains agents infectieux réussissent à franchir les barrières physiques, constituées, par exemple, par la peau, les muqueuses (app. respiratoire, app. génital, et app. digestif), l'épithélium à la surface des différents tissus, etc... Notre organisme essaie bien sûr de limiter la progression de ces agents pathogènes, et ceci dès les muqueuses, où quelques mécanismes de défense sont présents. Ces agents infectieux vont déclencher un ensemble de réactions de défense dans notre organisme, les plus virulents ayant développé des stratégies efficaces de contournement, voire de détournement du système de défense.

Un élément central de la protection de l'organisme, décrit dès les années 1882 par Metchnikoff, est la phagocytose (Peiser & Gordon, 2001), procédé par lequel il y a reconnaissance de l'élément pathogène, ce qui conduit à son internalisation par des cellules et à sa destruction (par dégradation). Ce procédé s'applique aux débris cellulaires, aux cellules apoptotiques, et, bien sûr, aux agents pathogènes présents dans le milieu extra-cellulaire. Ce procédé se déroule en synergie avec l'action du système du complément, que je détaillerai ci-dessous. La phagocytose et la cascade du complément constituent des éléments essentiels de la réaction immédiate à l'infection, en coopération avec l'action des neutrophiles, mastocytes, etc... Si l'intrus est éliminé rapidement, en quelques heures, l'infection passe inaperçue. Sinon, l'activation concertée de l'ensemble du système immunitaire se met en place, orchestrée par la production de diverses cytokines... Ceci permettra, si nécessaire, à la réponse immunitaire dite « adaptative » de prendre le relais au bout de quelques jours, temps nécessaire pour parvenir à une reconnaissance plus spécifique par les anticorps (produits par les lymphocytes B) et par les lymphocytes T. Il y aura ensuite mémorisation de cette reconnaissance spécifique, ce qui permettra une élimination très efficace des éléments pathogènes en cas de réinfection (d'où la notion d'immunité adaptative).

B.2. Présentation générale du système du complément.

Le terme « complément » n'a de sens qu'historique. En effet, à la fin du XIX^{ème} siècle, Jules Bordet découvrit l'existence d'un facteur thermo-sensible du sérum (appelé

initialement alexine) dont la présence était nécessaire pour observer l'activité bactéricide des anticorps, et même aussi leur activité hémolytique dans un système modèle utilisant du sérum d'animaux injectés de sang défibriné d'une autre espèce. Aujourd'hui, ce mot recouvre un ensemble d'au moins 30 protéines modulaires différentes (d'où l'appellation système), présentes dans la phase fluide et ou dans les membranes à la surface des cellules. Un litre de plasma contient plus de 3g de protéines du complément.

Le système du complément est l'une des armes les plus anciennes (d'un point de vue phylogénétique) de l'immunité innée et représente un composant essentiel de l'immunité « humorale » (dans le milieu extracellulaire). Il utilise un ensemble de protéines solubles qui réalisent une cascade de réactions de défense. Son action garantit en principe la limitation immédiate et efficace de l'infection, tout en préservant l'intégrité des cellules normales du « soi », qui sont protégées par la présence de protéines régulatrices à leur surface. Le système du complément a aussi un rôle important dans la modulation de la réponse immunitaire adaptative et dans l'élimination des cellules en cours d'apoptose et des complexes immuns circulants.

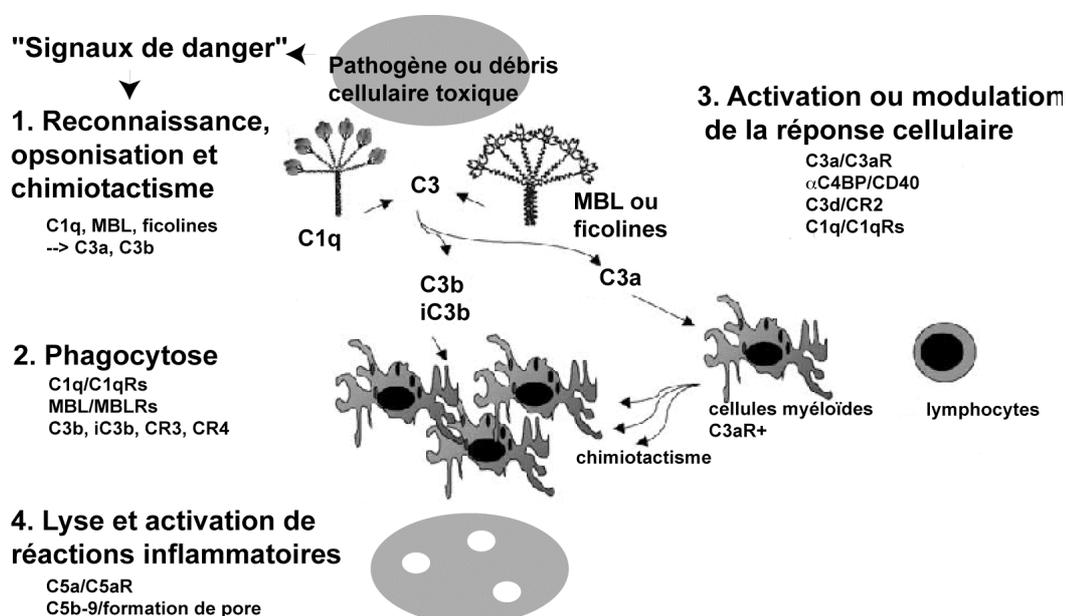


Figure B.2.1. Série de moyens de défense qui peuvent être activés par le système du complément. D'après Gasque, 2004.

Très schématiquement, le processus permettant d'éliminer un pathogène ou une cellule endommagée par le système du complément met en œuvre les étapes suivantes (Figure B.2.1): (1) les molécules de reconnaissance se fixent sur les motifs moléculaires signalant la surface cible et déclenchent une fixation covalente massive de fragments C3b (opsonisation), laquelle facilitera l'étape suivante de phagocytose (2). Les phagocytes sont attirés sur les lieux par les molécules chimiotactiques (C3a) produites simultanément. La production de ces fragments C3b et C3a met en jeu une cascade amplifiée de réactions protéolytiques analogue à celle de la cascade de la coagulation. La synthèse d'anaphylatoxines (fragments C3a, puis C5a) active et module la réponse cellulaire (3) et déclenche indirectement des signaux inflammatoires puissants. Ces derniers effets sont déclenchés par la fixation des protéines ou fragments du complément à des récepteurs cellulaires. Le système, s'il n'est pas limité par des protéines régulatrices, peut se poursuivre jusqu'à l'insertion finale d'un pore lytique dans la membrane reconnue (4), conduisant à la destruction du pathogène ou de la cellule infectée. De façon remarquable, les deux premières étapes

sont communes aux processus de défense immunitaire et d'homéostasie tissulaire (élimination douce des cellules apoptotiques et débris cellulaires). Il est maintenant bien établi que, dans le deuxième cas, la phagocytose ne déclenchera normalement pas de réaction inflammatoire et la réponse cellulaire sera modulée en direction de la tolérance.

B.2.1 La cascade protéolytique

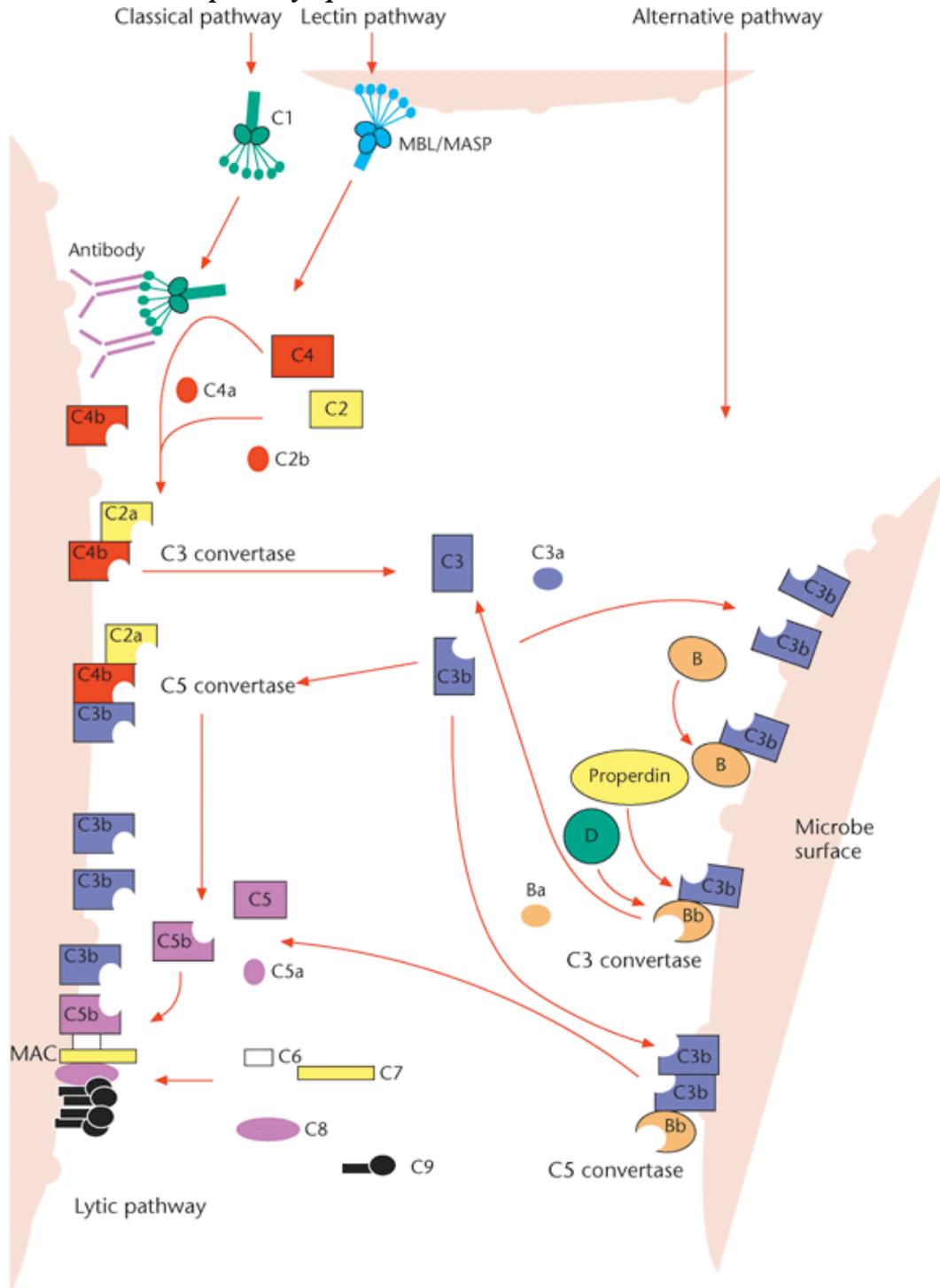


Figure B.2.2. Représentation schématique des cascades d'activation protéolytiques du système du complément. Les fragments C3b et C4b sont liés de façon covalente à la membrane par une liaison thioester. (figure extraite de Holland & Lambris, 2005)

B.2.2 Les voies d'activation

Le système du complément peut être activé par 3 voies distinctes, appelées les voies classique, lectine et alterne (figure B.2.2). L'activation par la voie alterne (impliquant les facteurs B, D, C3 et la properdine) se déroule au contact de « surfaces activatrices » des micro-organismes (décrites en 1988 comme riches en glycosylations dépourvues d'acide sialique : parois bactériennes, de levure, facteur du venin de cobra). Une molécule C3b se fixe sur cette surface, facilitant la fixation du facteur B qui peut être activé par le facteur D pour former le complexe C3bBb, qui pourra alors activer efficacement de nouvelles molécules C3. Les nouveaux fragments C3b, fixés à la surface, pourront à leur tour évoluer en complexes C3bBb, ce mécanisme d'amplification en boucle conduisant à un dépôt massif de fragments de C3 sur la surface ciblée.

Les voies classique et lectine sont activées par des complexes protéiques comprenant une protéine de reconnaissance multimérique à laquelle sont associées des protéases : C1q associé aux protéases C1r et C1s dans C1 pour la voie classique, Mannan Binding Lectin (MBL) ou ficoline associées aux MASPs (MBL-associated Serine Proteases) pour la voie lectine. La reconnaissance de plusieurs motifs moléculaires sur une surface va déclencher l'activation des protéases associées, lesquelles activeront ensuite par coupure protéolytique les facteurs C4 puis C2. Le complexe C4b2a (C3 convertase) ainsi créé permettra d'activer le facteur C3. Les motifs moléculaires reconnus sont souvent assimilables à des signaux de danger. Ils seront présentés de façon plus détaillée à la fin du chapitre. Les enzymes de la voie alterne peuvent là aussi former une boucle d'amplification à partir des fragments C3b, et donc conduire également à un dépôt massif de fragments C3b sur la surface reconnue.

B.2.3 La voie terminale

Les complexes C4b2a3b (voie classique ou lectine) et C3bBb3b (voie alterne) permettent la coupure de C5. L'étape ultime de ces cascades d'activation, dite voie terminale, implique les facteurs C5b, C6, C7, C8 et C9, dont l'assemblage produit le complexe d'attaque membranaire (MAC, appelé aussi le complexe C5b-9), qui peut s'insérer dans la double couche lipidique pour former un pore et induire la lyse de la cellule cible.

B.2.4. Système de contrôle.

Vu l'effet potentiellement dévastateur (lyse des cellules) du système du complément, un contrôle très strict de son action est absolument nécessaire (figure B.3.2). Un premier niveau de contrôle est assuré par le fait que l'action de la cascade protéolytique est localisée au niveau de la surface reconnue, les éléments C3 et C4 étant liés de façon covalente à cette surface. Ce contrôle est renforcé par un ensemble de régulateurs (appelés RCA pour Regulation of Complement Activation). Certains sont solubles (facteurs H et FHL-1, inhibiteur de C1, C4bp (C4 binding protein), facteur I, protéine S, clusterine ...). Les autres sont ancrés à la surface des cellules, dans les domaines enrichis en cholestérol (lipid rafts) : DAF (Decay Accelerating Factor or CD55), MCP (Membrane Cofactor Protein ou CD46), CD59 et CR1. Les facteurs solubles comme le facteur H ont la propriété d'adhérer aux membranes riches en acides sialiques, propriété caractéristique des cellules saines de l'organisme. Ces régulateurs empêchent ou limitent l'action du système du complément. Cette régulation se fait soit en inhibant l'action des protéases (C1r, C1s et MASPs), soit en

empêchant l'association du complexe d'attaque membranaire (C5b-9), ou des enzymes activant C3 et C5. Le facteur I, en présence de cofacteurs, clive les composants C3b en fragments inactifs iC3b, puis ensuite en fragments C3c, C3dg. Il clive aussi C4b.

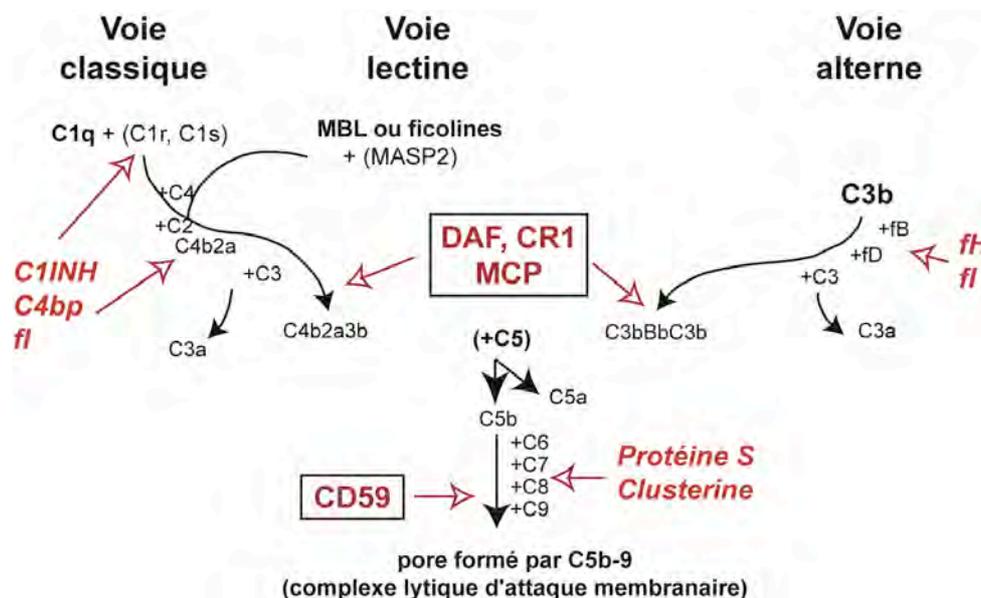


Figure B.3.2. Schéma très simplifié des points de régulation de la cascade du système du complément. Différents régulateurs (rouge) peuvent limiter la progression de la cascade enzymatique. Les régulateurs solubles, présents dans le milieu extracellulaire, sont en italiques, les régulateurs membranaires exprimés à la surface des cellules de l'hôte sont encadrés. D'après Gasque, 2004.

B.3 Protéines du complément : productions locales et autres fonctions

B.3.1. Tissus impliqués dans la synthèse des protéines du complément.

Le sérum plasmatique était considéré à l'origine comme la source unique de protéines du complément. En effet, chez les mammifères, l'essentiel (jusqu'à 90%) des protéines du complément est produit par le foie, à quelques exceptions près, comme C1q, le facteur D ou la ficoline M. Cependant, les protéines sériques ne peuvent pas atteindre librement certains tissus essentiels, comme le cerveau par exemple, protégé par la barrière hémato-encéphalique. Une production locale de protéines du complément a été identifiée depuis dans un grand nombre de tissus. En plus des hépatocytes, différents types de cellules synthétisent directement la plupart des composants du complément: monocytes, macrophages, cellules endothéliales ou épithéliales, fibroblastes, cellules dendritiques, cellules de la glie, astrocytes, pneumocytes, chorion, spermatozoïdes, cellules du ganglion de la rétine, etc... (Morgan et Gasque, 1997 ; Holland et Lambris, 2005). Cette production locale des protéines du complément est modulée par les cytokines. Elle peut être augmentée localement (et non dans le foie) au cours d'une phase de réponse aiguë par l'interféron gamma, les interleukines IL1 et IL6 et le TNF alpha. De plus en plus d'observations montrent que la sécrétion de protéines du complément (C3) par les cellules dendritiques conditionne leur capacité à activer la réponse primaire des lymphocytes T (Zhou et al, 2007).

B.3.2. Au-delà du rôle de protection immunitaire.

Même si les protéines du complément peuvent toujours jouer localement un rôle de défense immunitaire, quelques observations suggèrent ponctuellement un rôle qui dépasse largement le cadre strict de cette fonction, touchant les fonctions de reproduction, le développement, le métabolisme cellulaire ou l'homéostasie (Mastellos et al., 2005). Cela montre l'interconnexion des différents systèmes biologiques et suggère aussi finalement que ces protéines, apparues à des stades précoces de l'évolution, constituent la 'base' essentielle de certains processus vitaux, et non le 'complément' !. Je citerai ci-dessous quelques exemples impliquant entre autres la protéine C1q et ses récepteurs cellulaires:

- a) C1q est impliqué dans l'homéostasie tissulaire, contribuant à l'élimination rapide de cellules apoptotiques et d'autres formes altérées du « soi ».
- b) Les étapes précoces de l'activation du complément sont observées lors de la fusion acrosomiale, et pourraient être impliquées dans la phase d'entrée du spermatozoïde dans l'ovule (Riley-Vargas et al., 2005).
- c) Le développement anormal des tissus adipeux dans des cas d'obésité résistante à l'insuline semble lié à une activation excessive de la voie classique (Zhang et al., 2007)

B.4. Effets du système du complément relayés par les récepteurs cellulaires.

Plusieurs effets biologiques des facteurs du complément ou de leurs fragments sont la conséquence de leur interaction avec un ensemble de récepteurs cellulaires. Cette facette de l'activité du système du complément, qui permet par exemple une synergie d'action avec la phagocytose et le développement de la réponse des cellules immunitaires, est un domaine tout à fait passionnant mais les mécanismes moléculaires impliqués sont encore largement méconnus et en cours d'exploration. La recherche sur ce sujet n'est pas très aisée, car l'effet va dépendre de la distribution des récepteurs à la surface des différents types cellulaires (souvent variable en fonction des signaux produits par les cytokines environnantes ou du stade de différenciation de la cellule), des ligands reconnus, et de la coopération avec d'autres récepteurs. La plupart des phénomènes étudiés font en effet souvent intervenir la réticulation de différents récepteurs sur une surface ou l'association de différents récepteurs dans un contact étroit entre deux surfaces cellulaires, d'où la notion de 'synapses' phagocytiques (par exemple entre un macrophage et une cellule apoptotique) ou immunologiques (entre une cellule présentatrice d'antigène et un lymphocyte). Seuls quelques exemples seront rapidement évoqués.

C1q induit des effets variés par l'intermédiaire des récepteurs cellulaires du complément comme l'induction de chimiotactisme chez les neutrophiles et les mastocytes, la production des métabolites réactifs de l'oxygène (neutrophiles, lignée myéloïde) ou l'induction d'une réponse antiproliférative des lymphocytes B, avec diminution de la production d'IL1 et modulation de la production d'immunoglobines (Ghebrehiweit & Peerschke, 2004). D'autres effets modulateurs de la production de cytokines par les monocytes ont été récemment décrits pour C1q. Par exemple, son effet anti-inflammatoire serait corrélé à la production de complexes NFκB et à l'activation du facteur de transcription CREB (Fraser et al., 2007).

Le récepteurs des collectines cC1qR (ou calréticuline) et celui des têtes globulaires de C1q, gC1qR (p33), ne sont pas des protéines membranaires (Ghebrehiweit & Peerschke, 2004). L'analyse des mécanismes moléculaires impliqués nécessite donc d'identifier d'abord non seulement le(s) récepteu(s)r concerné(s) mais

aussi la(es) protéine membranaire adaptatrice, laquelle peut vraisemblablement varier en fonction des tissus et mécanismes considérés. Une exception est l'intégrine $\alpha 2\beta 1$, identifiée récemment comme récepteur de C1q (Edelson et al., 2006). Le rôle essentiel de cette interaction C1q/ $\alpha 2\beta 1$ vient d'être démontré *in vivo* dans le système modèle de la réponse à la *Listeria*. Cette réponse sollicite une partie importante du système immunitaire, et l'activation des mastocytes constitue un élément initial essentiel au système de défense. Les souris dépourvues d'intégrine $\alpha 2$ ont une déficience dans la défense précoce contre les bactéries (deux premiers jours), qui se multiplie beaucoup plus vite que dans les souches sauvages (Edelson et al., 2004). Cette réaction implique une coopération avec au moins un autre récepteur, qui reste à identifier (Edelson et al., 2006). L'autre exception est CR1.

CR1 est à la fois un récepteur et un régulateur du complément. CR1 reconnaît essentiellement les fragments opsonisés par C3b et C4b, mais aussi par C1q et MBL. Il est présent à la surface de toutes les cellules immunitaires, à l'exception de quelques sous-ensembles de lymphocytes T et cellules NK. Il est aussi absent de la surface des plaquettes. En présence de CR1 et de facteur I, les fragments C3b évoluent vers une forme C3dg. Ceci représente un changement important de stimulus : C3b porte un stimulus inflammatoire, alors que C3dg est potentiellement un stimulus régulateur des cellules B (par son interaction avec CR2). Vu le très grand nombre de globules rouges dans le sang, 85 à 90% des récepteurs CR1 sont situés à la surface de ceux-ci. C'est par cette voie que sont transportés par exemple les complexes immuns, jusqu'au foie et à la rate où ils seront éliminés par les phagocytes. Le récepteur CR1, à la surface des macrophages, déclenche la phagocytose des particules opsonisées. La présence de ce récepteur à la surface des cellules dendritiques folliculaires pourrait jouer un rôle dans le maintien de la mémoire immunologique, mais ce rôle est encore très controversé.

CR2 (appelé aussi CD21) est un récepteur ayant une affinité préférentielle pour les fragments C3d et C3dg. La forme longue de ce récepteur (16 modules CCP) est présente sur les cellules dendritiques folliculaires. Une forme plus courte (15 modules CCP) est présente à la surface des lymphocytes B, de cellules épithéliales et de quelques lymphocytes T. CR2 est présent de façon massive à la surface des lymphocytes B matures (8000 récepteurs /cellule !) alors qu'il n'est pas exprimé à la surface des lymphocytes B immatures ou pré-B, et qu'il disparaît lorsque le lymphocyte B se différencie pour sécréter les anticorps. Une altération très importante de la réponse primaire d'anticorps contre un antigène dépendant de la cellule T est observée en bloquant le récepteur CR2 chez la souris (anticorps anti-CR2, CR2 soluble, Souris CR2-). Le récepteur CR2 établit d'une part une interaction non covalente avec le complexe ternaire CD19/CD81(TAPA-1)/Leu13, d'autre part avec le C3dg à la surface de l'antigène reconnu par le récepteur BCR de l'antigène, établissant une « coligation » BCR-CD19 à la surface des lymphocytes B. Grâce à ce mécanisme de ligation, la liaison de C3dg à un antigène améliore la réponse immunitaire d'un facteur 10 000 comparé à l'antigène seul (Dempsey et al., 1996)

CR3 et CR4 sont les récepteurs principalement impliqués dans le déclenchement de la phagocytose. Ces récepteurs appartiennent à la famille des intégrines de type $\beta 2$, molécules d'adhésion hétérodimériques (CR3=CD11b/CD18 ; CR4=CD11c/CD18). Ils ont une affinité préférentielle pour les composants iC3b et C3b, ainsi que d'autres composants comme les lipopolysaccharides. CR3 est très répandu à la surface des monocytes (macrophages et cellules dendritiques), des cellules NK, et des neutrophiles. CR4 est certainement le récepteur de C3 le plus abondant sur les macrophages. Les mécanismes de phagocytose mis en jeu par ces récepteurs semblent différents de ceux déclenchés par la liaison aux récepteur Fc γ R,

les particules opsonisées semblant ‘sombrier’ rapidement dans la cellule, sans formation de larges pseudopodes (Peiser & Gordon, 2001). Ce mécanisme implique une réorganisation du cytosquelette, des microtubules et de l’actine, procédant via la protéine kinase Rho et non Rac comme dans le cas des récepteurs Fc γ R (Wiedemann et al., 2006). La phagocytose via CR3 dépend de la liaison de la taline au récepteur, ce qui n’est pas nécessaire dans le cas des récepteurs Fc γ R (Lim et al., 2007). Des observations récentes suggèrent également un rôle de ce récepteur dans le développement de l’immunité acquise. En effet, la ligation de ce récepteur sur les cellules dendritiques produirait un signal de ‘non danger’ en supprimant la synthèse de cytokines proinflammatoires, et donc en limitant la capacité des cellules dendritiques à stimuler et activer les lymphocytes T (Behrens et al., 2007). Plusieurs pathogènes ont adopté la stratégie de se fixer à ce récepteur.

Le récepteur CR1g est à rajouter à la liste des récepteurs phagocytiques des éléments opsonisés par les fragments de C3 (C3b et iC3b). Il est exprimé à la surface des macrophages résidant dans les tissus, en particulier dans le foie.

Les récepteurs des anaphylatoxines, C3aR et C5aR, sont des récepteurs transmembranaires à 7 hélices, couplés à des protéines G. Ces récepteurs sont impliqués dans de très nombreuses fonctions, dont la stimulation et le chimiotactisme des cellules de la lignée myéloïde, la dégranulation des mastocytes et leucocytes libérant les histamines, la modulation de signaux inflammatoires, l’adhésion. Ces récepteurs sont impliqués dans l’évolution de certaines situations pathologiques, comme la réponse allergique asthmatique ou le choc septique.

B.5. Système du complément et pathologies.

B.5.1. Pathologies liées aux déficiences en protéines du complément.

Les déficiences en composants du complément sont très rares.

Les déficiences compromettant les étapes d’opsonisation (C3b, iC3b) ou l’activité lytique (C6-C9) peuvent accroître très fortement la sensibilité aux infections bactériennes (streptocoques, staphylocoques). Les infections récurrentes par les bactéries du genre *Neisseria* sont également caractéristiques des déficits en facteurs terminaux. Une déficience en régulateurs du système du complément conduira plutôt à des pathologies montrant une altération non contrôlée de certains tissus de l’hôte: l’angioedème héréditaire, la glomérulonéphrite membranoproliférative, le syndrome hémolytique urémique (HUS) ou l’hémoglobinurie nocturne suraigüe (insertion du complexe lytique d’attaque membranaire dans la membrane des globules rouges).

De façon plus surprenante, les déficiences en protéines C1 et C4, se traduisent, en plus d’une sensibilité accrue aux infections bactériennes, par une prédisposition extrêmement forte (> 80%, >90% dans le cas de C1q) à un désordre autoimmun, le lupus érythémateux (SLE, systemic lupus erythematosus). Comment une déficience en ces éléments peut-elle aboutir à ce genre de désordre ? Ces protéines doivent avoir un effet protecteur par rapport au développement de ce type de maladie, qui peut survenir dans des circonstances diverses. Le mécanisme exact de cet effet protecteur n’est pas encore connu, et de nombreuses études visent à mieux comprendre le processus complexe de ce dérèglement, qui implique clairement le système du complément (Bao et Quigg, 2007). Une élimination déficiente de cellules apoptotiques (et de complexes immuns, très nombreux dans cette pathologie) intervient très probablement dans cette maladie. En effet, on a pu observer une accumulation de cellules apoptotiques dans le rein de souris C1q^{-/-} et C4^{-/-}, lesquelles exposent

vraisemblablement à leur surface des autoantigènes habituellement non accessibles. Une accumulation de cellules apoptotiques est également observée dans les centres germinatifs. Le mécanisme par lequel C1q pourrait jouer un rôle important pour l'établissement de la tolérance immunitaire n'est pas encore clairement identifié. Il a été montré cependant récemment que l'opsonisation de cellules apoptotiques par C1q facilite la phagocytose par les macrophages et par les cellules dendritiques, et que, de plus, cette opsonisation a des effets directs sur la modulation des messagers (interleukines et interféron) produits par la cellule dendritique après ingestion (Nauta et al., 2004) qui empêcherait le déclenchement de l'autoimmunité. Vue l'ampleur du dérèglement, il paraît vraisemblable que les récepteurs impliqués, en particulier CR1, jouent un rôle important dans le processus d'établissement de la tolérance immunitaire et que la progression de la maladie implique un déséquilibre croissant des signaux transmis par CR1, CR2, CR3, CR4, C3aR, C5aR, lesquels peuvent moduler leur niveau d'expression relatif (Bao et Quigg, 2007).

B.5.2. Dégâts liés à une activation excessive du système du complément.

Le système du complément est connu comme une arme à double tranchant, la balance entre ses effets positifs et potentiellement néfastes peut basculer dans certaines conditions pathologiques. En effet, les effets délétères du système du complément, et des réactions inflammatoires associées à son activation, peuvent aggraver certaines pathologies dès qu'il n'est plus correctement contrôlé.

On peut citer par exemple l'endommagement de tissus en cas de dépôt massif de complexes immuns, par exemple dans le rein (glomérulonéphrites). Dans le cas de brûlures non superficielles, l'activation du système du complément via les IgM, contre des auto-antigènes, déclenche une inflammation incontrôlée, ce qui va contribuer à aggraver la profondeur des lésions (Suber et al. 2007). Le système du complément semble impliqué dans la pathologie de la dégénérescence de la macula, altération irréversible de la vision centrale pouvant survenir chez les personnes âgées. Une activation du complément par la voie classique est aussi un caractère commun à de nombreux désordres neurodégénératifs, comme la maladie d'Alzheimer. Son implication dans l'évolution de la pathologie est donc probable. En effet, au-delà de son rôle protecteur, le système pourrait faciliter les agrégations fibrillaires en renforçant leur cytotoxicité, augmenter la sécrétion de cytotoxines inflammatoires et d'espèces réactives de l'oxygène, et enfin détruire des neurones par une production excessive du complexe lytique d'attaque membranaire (Bonifati et Kishore, 2007). De même, le système du complément semble effectivement jouer un rôle dans le processus de démyélination observé dans certaines maladies comme la sclérose en plaques (Urich et al., 2006 ; Jégou et al., 2007). Dans le cas des maladies de type prion, comme la maladie de Creutzfeld-Jacob, en plus des désordres neurodégénératifs, le système du complément pourrait également faciliter légèrement l'accumulation précoce des éléments toxiques PrPsc dans la rate, au niveau des cellules dendritiques folliculaires.

B.5.3. Modulation ou détournement du système du complément par certains pathogènes ou cellules cancéreuses.

Certains pathogènes réussissent à échapper au système du complément, voire à le détourner à leur avantage, profitant des systèmes de contrôle du complément protégeant l'hôte, ou du système particulier de phagocytose des cellules apoptotiques du soi, qui stimule la tolérance. Par exemple, les protéines de contrôle du complément, qui protègent les cellules de l'hôte, sont parfois copiées (vaccine) ou améliorées

(variole). Le virus VIH se protège directement avec les protéines régulatrices de l'hôte. Après activation du système du complément par C1q (liaison directe à la protéine gp41 ou aux anticorps), le virus sera recouvert de fragments de C3 progressivement désactivés (jusqu'à C3d), ce qui lui permettra d'adhérer successivement aux récepteurs CR1, sur les globules rouges, puis CR2, sur les lymphocytes B ou les cellules dendritiques, et d'être ensuite transmis efficacement aux lymphocytes T (Stoiber et al., 2005 ; Banki et al., 2006). L'infection des cellules dendritiques semble se faire via le récepteur CR3. Le virus d'Epstein-Barr utilise lui aussi le récepteur CR2 pour pénétrer dans les lymphocytes B et moduler la réponse immunitaire (Tanner et al., 1987). *Mycobacterium tuberculosis* utilise le récepteur CR4 pour pénétrer dans les macrophages au niveau des alvéoles pulmonaires, et modifie ensuite les protéines exprimées en surface de ces macrophages. Les pneumocoques utilisent plusieurs facteurs de virulence (capsule, PspA, PspC, pneumolysine) pour limiter l'action du système du complément, en fixant notamment le facteur H (PspC).

Les cellules cancéreuses sont aussi protégées par les régulateurs membranaires (CD46, CD55, CD59) ou solubles (clusterine) du complément ce qui va limiter son action cytolitique dans les stratégies de vaccination contre les tumeurs (Fishelson et al., 2004), à moins d'utiliser des anticorps qui neutralisent également ces régulateurs (Ziller et al., 2005).

B.6. Informations structurales sur les protéines du complément.

Encore timide à la fin des années 1990, la biologie structurale des protéines du complément a pris son essor dans les années 2000, et maintenant une petite centaine de structures sont déposées dans la PDB, avec une certaine redondance, bien sûr. Nous avons ainsi déposé 20 structures, qui seront brièvement présentées et analysées dans les chapitres C et D. Le tableau suivant montre les principaux fragments dont la structure a été résolue dans les autres équipes. Dans un premier temps, seules les structures de petites protéines (facteur D) ou de fragments, de modules isolés ou paires de modules ont été définies. Récemment, les structures de protéases modulaires entières (facteurs B et C2) et celle du fragment central C3 (1641 a.a.) ont été résolues.

Différents niveaux d'informations ont pu être obtenus :

- a) la définition du repliement, lorsque celui-ci n'était pas extrapolable à partir d'autres structures connues. Ex : C3. Dans ce cas, même le découpage en modules n'avait pas été correctement anticipé à partir de l'analyse des séquences.
- b) La position relative des différents modules ou domaines, et leur variation au cours des séquences
- c) La définition précise de la structure des boucles de surface, variables d'une protéine à l'autre. Cela constitue un support important pour la localisation éventuelle des informations de mutations pathologiques, de sites fonctionnels identifiés par mutagenèse dirigée
- d) Les degrés de liberté de certains mouvements aux charnières intermodulaires.
- e) La visualisation d'interactions (récepteur ligand, assemblage de complexes)

Les structures des protéases B, D et C2 ont apporté des informations très intéressantes sur les mécanismes de régulation sophistiqués de leur activation. La structure de la protéine centrale C3 et celle de ses différents fragments permettent

d'observer des déplacements très importants de blocs modulaires après les premières étapes de clivage enzymatique de cette protéine et donc de mieux visualiser les changements de signaux correspondant à ses clivages successifs. Les études structurales par résonance magnétique nucléaire (RMN) ont fourni des informations importantes sur les sites de fixation de différents régulateurs, et sur les éléments de flexibilité qui modulent leur orientation relative. Des informations structurales sur certains fragments de récepteurs sont également disponibles.

Complement protein	PDB code	Exp Meth	Reference
Human Factor D (active/ proenzyme))	1DSU 1FDP	X-ray	Narayana, et al (1994) J. Mol. Biol. 235, 695-708 Jing, et al (1999) Embo J. 18 804-814
Human Factor B Factor B	1DLE 2OK5	X-ray	Jing, et al (2000) Embo J. 19 164-173 Milder et al., (2007) Nat. Struct. Mol. Biol. 14, 224-
Human MASP-2: fragment CCP2-SP	1Q3X	X-ray	Harmat, et al (2004) J. Mol. Biol 342 1533-1546
Rat MASP-2: fragment CUB1-EGF-CUB2	1NT0	X-ray	Feinberg, et al (2003) EMBO J. 22 2348-2359
Human C4dg	1HZF	X-ray	Van Den Elsen, et al (2002) J. Mol. Biol. 322 1103.
Human C2a	2I6Q	X-ray	Milder et al., (2006) Structure, 14, 1587-97
Human C3d C3dg C3c C3 C3 bovine C3b C3b + CRiG	1C3D 1QSJ 2A74 2A73 2B39 2HR0 2IO7 2ICF	X-ray	Nagar, et al (1998) Science 280, 1277-1281 Zanotti et al. (2000) BBA, 1478, 232-8 Janssen et al. (2005) Nature, 437, 505-511 Fredslund et al. (2006) J. Mol. Biol. 361, 115-127 Abdul Ajees et al. (2006) Nature, 444, 221-225 Janssen et al. . (2006) Nature, 444,213-216 Wiesmann et al. . (2006) Nature, 444, 217-220
hC3d/CR2(CCP1-2)	1GHQ	X-ray	Szakonyi, et al (2001) Science 292, 1725-1728
Human C5a C345C	1KJS 1XWE	NMR	Zhang, et al (1997) Proteins: 28 261-267. Bramham (2005) J Biol Chem., 280, 10636-45
Human C8 γ	1IW2	X-ray	Ortlund, et al (2002) Biochemistry 41, 7030-7037.
hCR1 CCP 15 -16.	1GKG	RMN	Smith, et al. (2002) Cell 108 769-780
hCR2 CCP1-2	1LY2	X-ray	Prota, et al (2002) PNAS, 99, 10641-10646
hCR3a (CD11b)	1IDO	X-ray	Lee, et al (1995) Cell 80 631-638
hCR3_ β 2_EGF (CD18) _I-EGF	1L3Y 1YUK	NMR X-ray	Beglova, et al (2002) Nat. Struct. Biol. 9, 282-287 Shi et al. (2005) J Biol Chem., 280, 30586-93
gC1qR	1p32	X-ray	Jiang et al. (1999) PNAS, 96, 3572-3577.
α 2 β 1 I domain	1QC5	X-ray	Rich et al. (1999) J. Biol. Chem., 274, 24906-13
CD55 (DAF) CCP1-4 CCP3-4	1OJV 1OH3	X-ray	Lukacik, et al (2004) PNAS, 101, 1279-1284 Williams et al. (2003) J Biol Chem. 278, 10691-6.
CD59 (1-70) extracell.	1ERG	NMR	Kieffer, et al (1994) Biochemistry 33, 4471-4482
Factor H CCP15-16 CCP19-20 CCP7	1HFH 2BZM 2JGW	NMR	Barlow et al. (1993) J Mol Biol. 232, 268-284 Herbert et al. (2006) J. Biol. Chem., 281, 16512-20 Herbert et al. (2007) J. Biol. Chem. In press
C4BP α CCP1-2	2A55	NMR	Jenkins et al., (2006) J Biol Chem, 281, 3690-7.
CD46 CCP 1-2 (+ adenovirus 11 knob)	1CKL 2O39	X-ray	Casasnovas et al., (1999) EMBO J. ,18, 2911-2922 Persson et al. (2007) Nat Struct Mol Biol.,14, 164-6.
BBCRASP <i>Borrelia Burgdorferi</i>	1W33	X-ray	Cordes, et al.(2005) Nat Struct Mol Biol. 12, 276-7
Vaccinia VCP (+ héparine)	1G40 1RID	X-ray	Murthy, et al. (2001) Cell 104 301-311. Ganesh, et al. (2004) PNAS 101 8924-8929.
C1 inhibitor	2OAY	X-ray	Beinrohr et al., (2007) J. Biol. Chem., in the press

B.7 Les protagonistes de l'activation spécifique du complément.

Les sujets sur lesquels j'ai travaillé ces dernières années concernent plus spécifiquement l'étude structurale des étapes de reconnaissance et d'activation par les complexes protéiques permettant l'initiation des voies classique et lectine du complément. Je vais donc détailler un peu plus les molécules impliquées et leurs fonctions respectives dans ce processus. Ces complexes comprennent une protéine de reconnaissance à laquelle sont associées des protéases (Figure B.7.1).

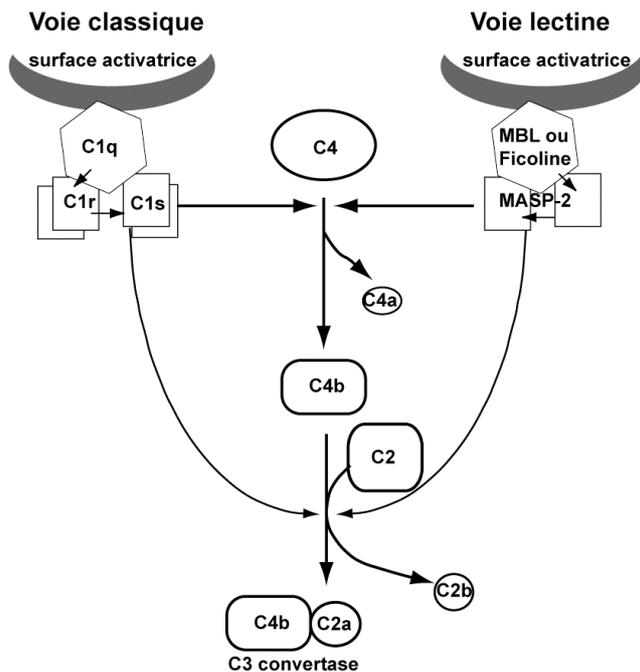


Figure B.7.1. Schéma des cascades d'activation moléculaire du système du complément par les voies classique et lectine.

Les protéines de reconnaissance sont marquées par des hexagones ; les protéases associées par des carrés. Les flèches marquent les processus d'activation successifs. Les protéases C1s et MASP2 clivent d'abord C4 puis C2. Les deux voies aboutissent à la formation du même complexe C4bC2a ou « C3 convertase », capable de cliver C3, élément central du système du complément.

B.7.1. Le complexe C1 et l'activation de la voie classique.

Le complexe C1 se compose de la protéine de reconnaissance multimérique C1q à laquelle s'associe un tétramère (dépendant du calcium) de deux protéases C1r et C1s. L'émergence de ces protéases au cours de l'évolution semblent coïncider avec le développement initial du système de défense adaptatif (Figure B.7.2). La fixation multivalente du complexe C1 sur une surface activatrice (agent pathogène, cellule apoptotique, ...) se fait par l'intermédiaire de la protéine C1q. Le signal fourni par cette étape de fixation va déclencher l'activation des protéases C1r associées, lesquelles activeront à leur tour les protéases C1s, qui activeront la cascade du complément par coupure des composants C4 puis C2.

B.7.2. Les complexes MBL-MASPs et ficoline-MASPs de la voie lectine .

L'activation du complément par la voie lectine a été découverte beaucoup plus récemment, sans doute parce que les MASPs et les ficolines sont en faible concentration dans le sérum. Cependant cette voie d'activation semble de loin la plus ancienne du point de vue de l'évolution, retrouvée chez quelques invertébrés assez évolués comme les ascidies (prochordés) et des vertébrés très primitifs comme la lamproie, ce qui suggère que ce système est apparu depuis plus de 600 millions d'années (Figure B.7.2).

La MBL (Mannan Binding Lectin) et les ficolines jouent le même rôle de reconnaissance et d'activation que la protéine C1q dans la voie classique. Comme C1r

et C1s, la protéase associée MASP2 permet la coupure des composants C4 et C2. Cependant, il existe 3 protéases MASPs différentes. Seule la protéase MASP2 est clairement capable de couper les substrats C4 puis C2, et même de s'autoactiver (remplaçant ainsi à la fois C1r et C1s). Le rôle des autres protéases MASP1 et MASP3, qui sont deux produits de l'épissage alternatif du même gène MASP1, est encore à éclaircir. De plus, le contenu des complexes MBL-MASPs apparaît variable, la variabilité portant sur le nombre de chaînes de MBL et la nature des MASPs associées (MASP2, MASP1, MASP3 ou MASP19).

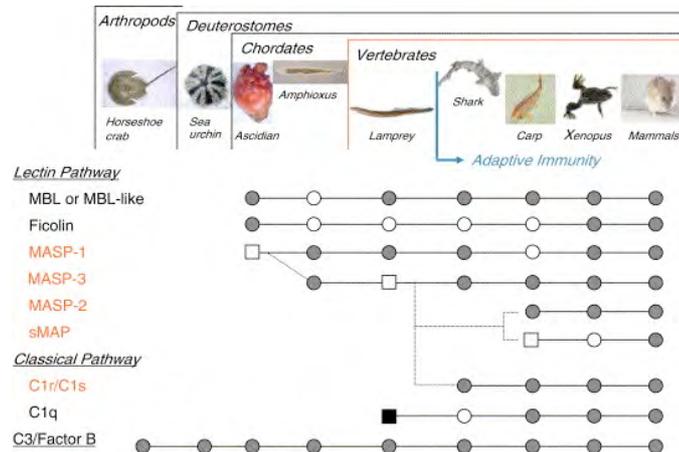


Figure B.7.2. Apparition des molécules activant les voies de type lectine puis classique au cours de l'évolution, précédant celles du système adaptatif. Les ronds pleins marquent l'identification de protéines similaires, les carrés indiquant des différences de propriétés biochimiques plus marquées. (Endo et al., 2006)

B.8. Les protéines de reconnaissance : C1q, MBL et ficolines.

Toutes ces protéines possèdent de longues tiges de type triple-hélice de collagène suivies par un domaine de reconnaissance de « signal de danger » en C-terminal, d'où l'appellation générique de « collagènes de défense », qui n'est cependant pas utilisée de façon générale.

B. 8.1. C1q.

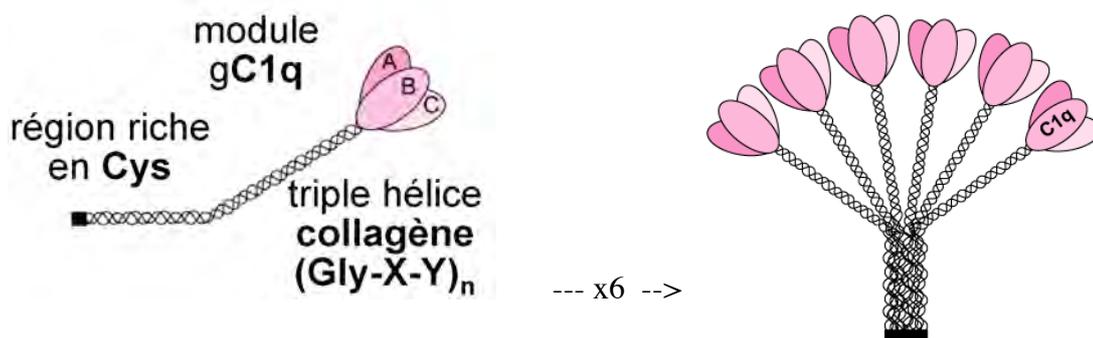


Figure B.8.1. Structure modulaire de C1q

Trois sous-unités de C1q (A : 221 a.a., B : 227 a.a., C : 217 a.a.) s'associent en hétérotrimère formé d'une triple hélice en N-terminal et d'un domaine globulaire en C-terminal (modules gC1q). Comme le suggère l'image de microscopie électronique (Knobel et al., 1975), six hétérotrimères s'associent en forme de bouquet de tulipe.

La sous-unité C1q comprend 18 chaînes polypeptidiques de trois types différents (A, B, C), chacune d'une longueur d'environ 200 acides aminés et de séquence homologue (Reid, 1979). Des analyses de C1q par microscopie électronique (Figure B.8.1) ont montré une structure oligomérique ressemblant à un bouquet de six tulipes, où les fleurs représentent les têtes globulaires (Knobel et al., 1975). Les régions N-terminales des six hétérotrimères forment une tige. Du fait de l'interruption dans la séquence de type collagène Gly-X-Y, les triples hélices divergent pour former six « bras », divergeant à partir d'une charnière flexible. Vers l'extrémité C-terminale de chaque hétérotrimère, les trois chaînes polypeptidiques se séparent pour former trois modules de type gC1q qui, ensemble, forment la tête globulaire de C1q.

B. 8.2. MBL et ficolines

L'organisation de la protéine MBL est similaire, mais le domaine gC1q C-terminal est remplacé par un domaine lectine de type C (Figure B.8.2). Des images de microscopie électronique de MBL de rat (Lu et al., 1990) ont montré le même genre de structure macroscopique que C1q, avec un assemblage pouvant comprendre de 3 à 6 trimères, 3 et 4 étant les plus fréquemment observées (publi n° 12). La MBL possède en plus un domaine de type 'coil-coil' (région 'cou') précédant les domaines CRD, et nécessaire à la formation correcte du domaine de reconnaissance.

Dans le cas des ficolines, le domaine C-terminal est composé de trois modules de type fibrinogène. D'autre part, les tiges de collagène semblent partir d'un domaine commun en N-terminal ayant une forme plus globulaire (Figure B.8.2).

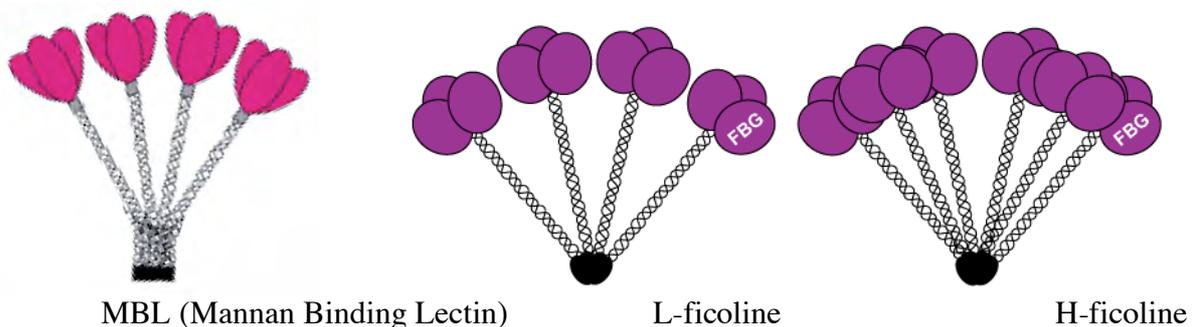


Figure B.8.2. Structures modulaires de la MBL et des ficolines.

Par analogie avec C1q, la même représentation sous forme de « bouquets » a été adoptée pour la MBL, même si les images de microscopie électronique montrent moins clairement le « tronc » commun de départ. Les domaines globulaires sont ici des homotrimères, avec des domaines de type lectine C pour la MBL, et de type fibrinogène pour les ficolines

B.8.3. Principe de reconnaissance de motifs par avidité collective.

La structure du domaine lectine de type C de la MBL de rat a été la première déterminée dans cette sous-famille de protéines (Weis et al., 1992). Cette structure a permis de déchiffrer les déterminants structuraux permettant la reconnaissance des deux hydroxyles en configuration équatoriale sur le mannose. Cependant, cette reconnaissance d'un micromotif est caractérisée par une affinité beaucoup plus faible ($K_d > 10^{-3}$ M) que celle obtenue par la fixation de la MBL entière (K_d de l'ordre de 10^{-9} M) sur une surface présentant ce motif répété selon une géométrie qui permet la reconnaissance simultanée par plusieurs domaines (Figure B.8.3). Cette fixation collective conduit à un phénomène d'avidité pour la surface reconnue, permettant d'obtenir ainsi une reconnaissance sélective très efficace à partir de sites de fixation d'affinité modérée. Ce mode de fonctionnement est valable pour l'ensemble des

'collagènes de défense'. La fixation de ces molécules implique donc deux niveaux hiérarchiques de reconnaissance : celle de chaque micromotif, puis celui de leur répartition spatiale, définissant le macromotif (Figure B.8.3).

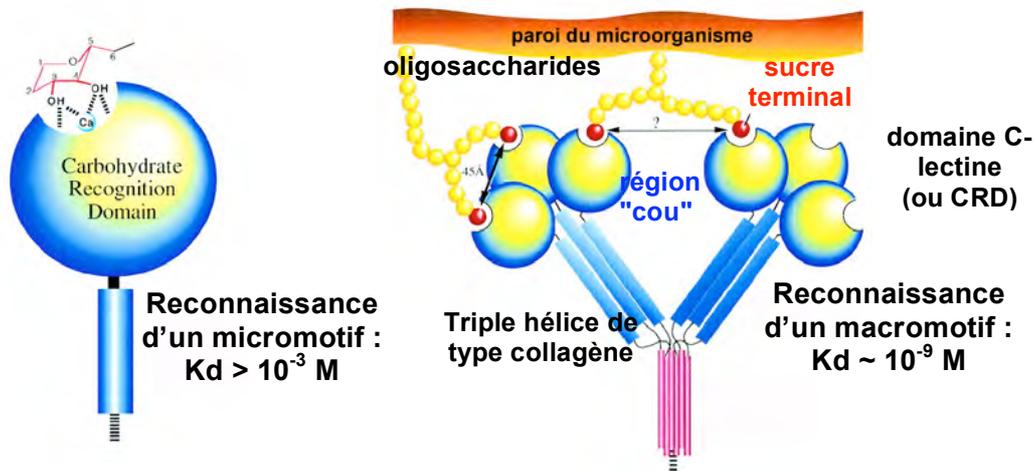


Figure B.8.3. Illustration de l'avidité de fixation de la MBL sur une surface.

B.9. Structure modulaire commune des protéases associées : C1r, C1s et MASPs.

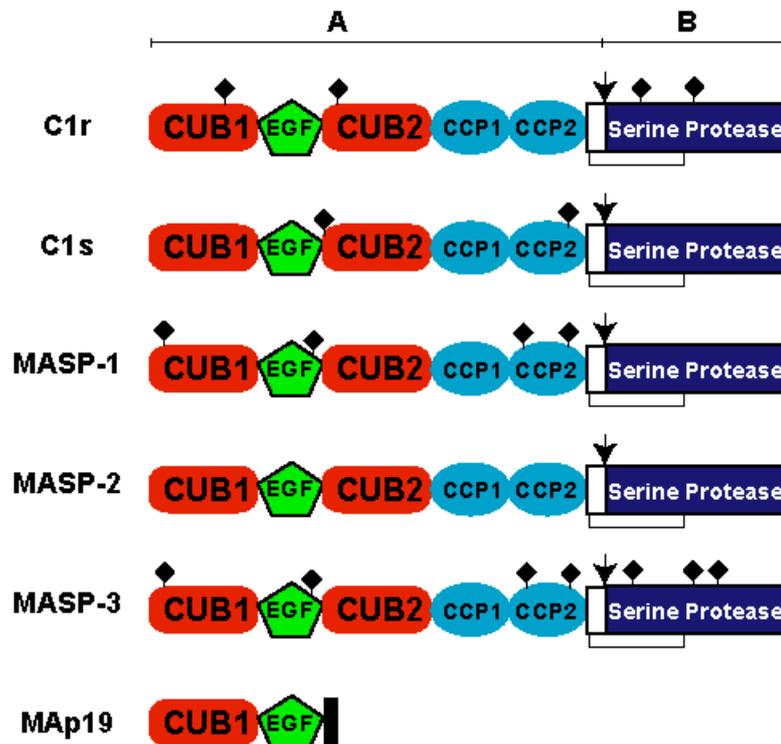


Figure B.9. Structures modulaires des protéases C1r, C1s, MASPs et de la Map19. Les flèches indiquent la position de la liaison Arg-Ile qui est coupée au cours de l'activation des protéases, donnant deux chaînes A et B liée par un pont dissulfure. ♦ indique un site de glycosylation. ■ représente l'extension C-terminale EQSL de la Map19.

Les protéases associées à C1q, aux ficolines ou à la MBL présentent une structure modulaire commune (Figure B.9). L'activité catalytique est portée par le domaine protéase à sérine, en C-terminal, précédé de deux modules CCP. La partie N-terminale, constituée de deux modules CUB encadrant un module de type EGF (Epidermal Growth Factor), permet l'association aux protéines de reconnaissance, ainsi que l'association des protéases entre elles : interaction entre C1r et C1s et formation d'homodimères pour les MASps. Toutes ces associations dépendent de la présence de calcium.

L'appellation de module CUB provient des initiales des 3 protéines dans lesquelles ce module a été initialement mis en évidence : C1r/C1s, Uegf, et BMP (bone morphogenetic protein). Ces modules sont trouvés dans des protéines supportant des fonctions diverses, dont le système du complément et les processus de développement tels la fertilisation, l'embryogénèse et l'organogénèse. Les modules CUB comportent environ 110 acides aminés, dont 4 cytéines formant deux ponts disulfures, sauf dans les modules CUB de C1s, C1r et MASPs qui ne possèdent pas le premier pont disulfure.

Les modules de type CCP (Complement Control Protein) sont trouvés presque exclusivement dans des protéines du complément, en particulier dans des protéines de contrôle du complément (RCA), y compris dans le récepteur CR1 ou dans les protéines de contrôle issus des micro-organismes (par exemple la protéine de la vaccine VCP). Ces modules comportent environ 60 à 79 acides aminés. Il y a également 4 cystéines, formant deux ponts disulfures.

MASP1 et MASP3 d'une part, MASP2 et Map19 d'autre part, sont des produits par l'épissage alternatif d'un même gène. MASP1 et MASP3 ne diffèrent que par le domaine protéase à sérine.

B.10. Marqueurs moléculaires reconnus par C1q et les ficolines.

Le dernier partenaire à prendre en compte dans le système d'activation est le(s) motif(s) moléculaire(s) reconnu spécifiquement par les protéines de reconnaissance, ayant un rôle physiologique clair '*in vivo*'. Celui-ci est souvent très délicat à identifier, et, dans le cas de la molécule C1q, il peut ne pas être unique. Un répertoire surprenant de cibles diverses a été mis en évidence pour la protéine C1q (publi n° 3). Pour les ficolines, les études ont été plus limitées, mais la détermination d'une spécificité de reconnaissance pour des sucres simples (par liaison directe ou compétition) est plus facile.

B.10.1 C1q.

La liste ci-dessous vise surtout à illustrer la grande diversité des ligands identifiés pour la molécule C1q et l'hétérogénéité des informations les concernant. Au-delà d'une attraction à distance pour des groupes de charges négatives, il paraît impossible de définir quelques marqueurs moléculaires communs à ces différentes cibles !

◇ **Anticorps.** La protéine C1q est connue essentiellement pour sa reconnaissance spécifique des complexes anticorps-antigènes. Ces complexes, fixés en surface, indiquent la présence d'un élément étranger reconnu par un anticorps (IgG ou IgM). Les complexes immuns circulant dans le sang sont des éléments potentiellement nocifs, à éliminer.

◇ **Protéines de la famille des pentraxines :** SAP (Serum Amyloid P component), CRP (C reactive Protein) et la pentraxine 3 (PTX3). La fixation de ces protéines marque souvent la présence de cellules endommagées, de microorganismes, de cellules apoptotiques...

◇ **le récepteur cellulaire DC-sign R1.** Cette interaction a été identifiée récemment dans le cadre d'une étude sur les éléments de défense essentiels contre les pneumocoques dans la

rate (Kang et al., 2006). L'activation de la voie classique après reconnaissance par C1q des molécules SIGN-R1 fixant la capsule des bactéries a été proposée justement comme un élément essentiel de notre défense innée contre ces pathogènes.

◇ **Cellules en cours d'apoptose.** Cette opsonisation facilitera la phagocytose de ces cellules par les macrophages et aussi par les cellules dendritiques. Le(s) marqueur moléculaire reconnu n'est pas clairement défini. L'hypothèse d'une reconnaissance directe de la glycérophosphosérine a été explorée par mes collaborateurs à l'IBS. D'autres groupes suggèrent une reconnaissance de la calréticuline, ou une fixation indirecte via des pentraxines ou des IgM.

◇ **Porines de certaines bactéries,** comme celles de *Klebsiella pneumoniae* (OMP36), d'*Aeromonas salmonicida*, de *Mesophilic Aeromonas hydrophila* (porin II), de *Legionella pneumophila* (ompS), de *Salmonella minnesota*. C1q pourrait inhiber ainsi la formation de pores et activer le système du complément. Les lipopolysaccharides sont aussi reconnus par C1q, au moins pour *Mesophilic Aeromonas hydrophila*.

◇ ***Streptococcus pyogenes*** et d'autres streptocoques du groupe A (Bactérie gram+). Liaison directe aux protéines M5 et FcRA76. Cependant les protéines M de cette bactérie peuvent fixer le fibrinogène ou la protéine régulatrice C4bp et ainsi protéger d'une attaque par le système du complément.

◇ ***Mycoplasma pneumoniae*** (molécule non identifiée)

◇ **Certaines protéines d'enveloppe de rétrovirus** peuvent aussi être reconnues : gp41 et gp120 de HIV-1, gp21 de HTLV-1, p15E de MuLV. Tous les rétrovirus activent le système du complément mais ils ont développé différents moyens pour limiter la lyse du virus, profitant de l'opsonisation pour mieux adhérer aux récepteurs du complément.

◇ PrP_{scr}, forme infectieuse du prion PrP.

◇ les peptides N-terminaux des peptides β amyloïde de la maladie d'Alzheimer et d'autres formes de démence familiale.

◇ **Protéines de la myéline :** MOG (Myelin Oligodendrocyte-Glycoprotein) et MBP (Myelin Basic Protein),

◇ l'**héparine** et autres héparanes sulfates

◇ **Protéines d'absorption:** human salivary scavenger protein gp340. Site non défini.

◇ **polyanions** (ADN, ARN, ...).

◇ **Des protéines glycosylées de la matrice extracellulaire :** fibromoduline, décorine et 'biglycan'. La décorine aurait un rôle inhibiteur de l'activation de C1 et une fixation vraisemblable sur les tiges de type collagène, se liant aussi à d'autres collectines. La fibronectine se lie à C1q. C1q se lie à la surface des plaquettes et des cellules endothéliales.

B.10.2. Ficolines

Les spécificités moléculaires des ficolines seront reprises plus loin sur quelques exemples, leur point commun étant de reconnaître les composés acétylés. Les cibles principales des ficolines sont les agents pathogènes bactériens et les cellules apoptotiques (ou nécrotiques). La ficoline L reconnaît l'acide lipotéichoïque et la capsule de plusieurs bactéries à gram positif (Lynch et al., 2004 ; Krarup et al., 2005), mais aussi un glucane fongique (Ma et al., 2004). La ficoline H reconnaît un répertoire beaucoup plus restreint de bactéries à gram positif (essentiellement *Aerococcus viridans*) et les acides lipopolysaccharidiques de Salmonelles (*S. typhimurium* et *S. minnesota*). La ficoline α de porc reconnaît les acides lipotéichoïques de quelques bactéries à gram positif et les acides lipopolysaccharidiques de nombreuses bactéries à gram négatif (Nahid et Sugii, 2006). La ficoline M n'est pas encore bien connue, mais a pour particularité de fixer les acides sialiques.

B.11. Objectifs de l'étude structurale.

Je vais maintenant présenter l'étude structurale de différents fragments fonctionnels de complexes permettant l'initiation 'spécifique' des voies classique ou lectine du complément. En plus de la structure isolée de ces modules et domaines, nous serons intéressés par l'observation des interactions entre ces éléments (formations de dimères par exemple), les interactions avec des ligands (pour les protéines de reconnaissance). Nous verrons aussi comment ces études peuvent aider à répondre à la question fondamentale suivante : **Comment ces molécules sont-elles organisées pour que le processus de reconnaissance de signaux moléculaires, effectué par les têtes globulaires, déclenche de façon contrôlée l'activation des protéases associées, lesquelles activeront ensuite spécifiquement la cascade d'effets biologiques induits par le système du complément ?**

C. Etudes structurales du complexe C1.

L'analyse structurale du complexe C1 est rendue difficile de par sa taille (790 kDa), sa flexibilité et la variété de ses états d'activation. Une stratégie de dissection fonctionnelle a été mise en place, ce qui a permis de progresser malgré tout dans la connaissance des relations structure-fonction de cet assemblage complexe. Je résumerai cette approche ainsi que les principaux résultats qui permettent de mieux comprendre les mécanismes moléculaires impliqués dans le fonctionnement de ce complexe, qui vont de la reconnaissance d'un « signal de danger » au déclenchement induit d'une cascade de réactions enzymatiques

C. 1. Dissection et reconstitution du complexe C1 .

C.1.1. Stratégie de dissection structurale et fonctionnelle.

Nous avons choisi de limiter dans un premier temps les études structurales à des fragments fonctionnels stables et homogènes. Ceux-ci ont d'abord été obtenus dans le groupe de Gérard Arlaud en utilisant la technique de protéolyse ménagée. Des fragments recombinants ont ensuite été produits pour les protéases C1r et C1s, cette technique étant encore difficile à mettre en œuvre dans le cas plus complexe de la protéine C1q. L'utilisation de la biologie moléculaire, qui permet la production de mutants et/ou de fragments de différentes longueurs et leur analyse fonctionnelle et cristallographique, a largement contribué au succès de cette approche.

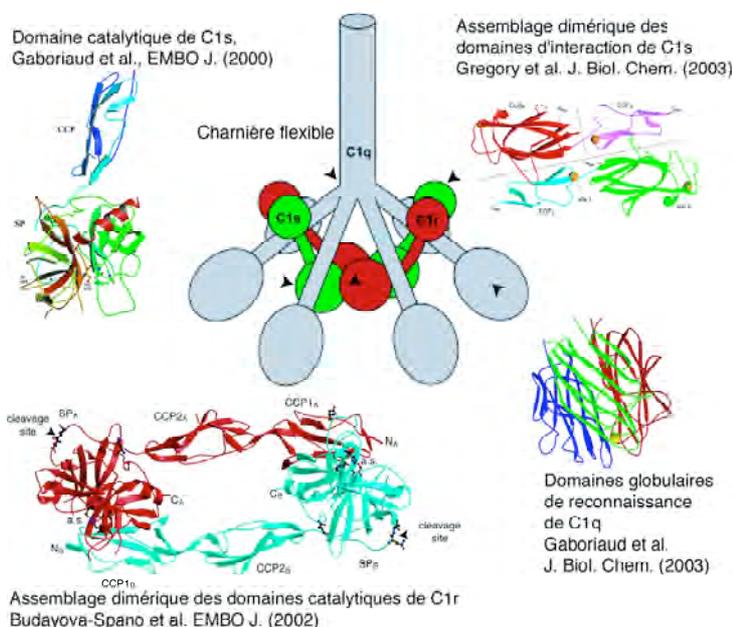


Figure C.1.1.
Modèle fonctionnel initial du complexe C1 (d'après Arlaud et al., 1987) et position dans ce contexte des domaines fonctionnels dont nous avons résolu la structure cristallographique.

C.1.2. Choix des fragments fonctionnels.

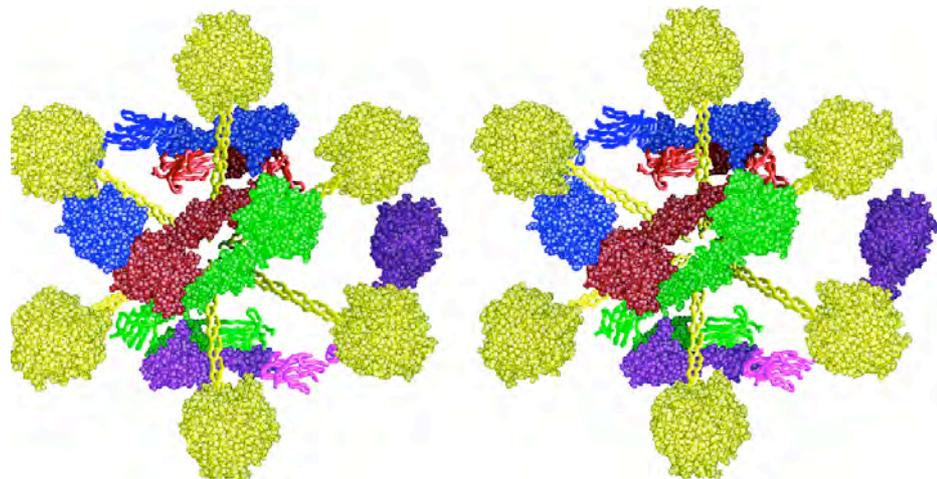
Nous nous sommes intéressés en priorité aux fonctions de reconnaissance (portée par le domaine globulaire de C1q) et d'activation (portée par les domaines catalytiques des protéases C1r et C1s). Pour mieux comprendre les éléments structuraux permettant l'association du tétramère C1s-C1r-C1r-C1s et son association à C1q, nous avons également entrepris l'étude structurale du domaine d'association de la protéase C1s. L'ensemble de ces structures nous a aidé à mieux comprendre le mécanisme de transmission de signal entre la protéine C1q et les protéases associées.

C.1.2. Affinement du modèle macroscopique du complexe C1.

Comme le montre la figure C.1.1., un modèle macroscopique et fonctionnel de C1 avait été proposé (Arlaud et al., 1987). Celui-ci intégrait les informations obtenues à l'époque par des méthodes variées comprenant le découpage en domaines fonctionnels (protéolyse ménagée) et leur analyse par biochimie, microscopie électronique, diffusion de neutrons,... La figure C.1.1 permet de situer, dans le contexte de ce modèle initial, les différents fragments fonctionnels dont nous avons résolu la structure.

Nous avons pu intégrer l'ensemble de ces structures (qui représentent les 2/3 de la structure de C1) dans un modèle plus précis (Figure C.1.2), où les fragments dont nous n'avons pas déterminé la structure ont été remplacés par des fragments homologues ou des modèles (pour les tiges de type collagène). L'intérêt de ce modèle affiné est qu'il permet de représenter tous les éléments à la même échelle (atomique), de tester la cohérence de cet assemblage final et de proposer des hypothèses quant au mécanisme d'activation du complexe (publis n°2, 5, 7 et 35), lesquelles seront reprises et explicitées en fin du chapitre C (C.6).

Figure C.1.2.
Modèle affiné
du complexe
C1. (vue de
 dessous
 stéréoscopique).
 C1q, jaune ;
 C1rA : rouge,
 C1rB : vert,
 C1sA : bleu ;
 C1sB : magenta.
 Sphères pleines :
 structures expéri-
 mentales.
 rubans: modèles
 homologues



C.1.3. Le « fil rouge » de la présentation concise des résultats.

Les différentes études structurales, qui ont permis d'arriver à ce résultat global, vont être reprises brièvement. Tous ces résultats étant publiés, le résumé sera focalisé sur les éléments essentiels qui aident à mieux comprendre les relations entre structure et fonction pour chacun de ces domaines fonctionnels. Les nombreuses comparaisons de structures et

de séquences avec les fragments ou modules homologues seront omises par souci de concision. Les détails expérimentaux ne seront pas non plus repris dans cette partie. Enfin, le résumé des études structurales ne sera pas réalisé en suivant l'ordre chronologique de leur résolution, mais suivra plutôt l'ordre de la séquence d'évènements permettant la transmission du signal d'activation.

C.2. Le domaine globulaire de reconnaissance de la protéine C1q.

L'évènement initial du processus d'activation est la reconnaissance simultanée, par plusieurs têtes globulaires de C1q, d'un signal de « danger ». C1q semble capable de reconnaître une très grande variété de signaux (cf paragraphe B.2.2), même si le plus connu et le plus étudié correspond à sa fixation sur des complexes immuns. D'où la question de mieux voir et comprendre quelles bases structurales permettent d'atteindre une si grande variété de détection.

C.2.1. Structure d'une « tête globulaire » trimérique de C1q.

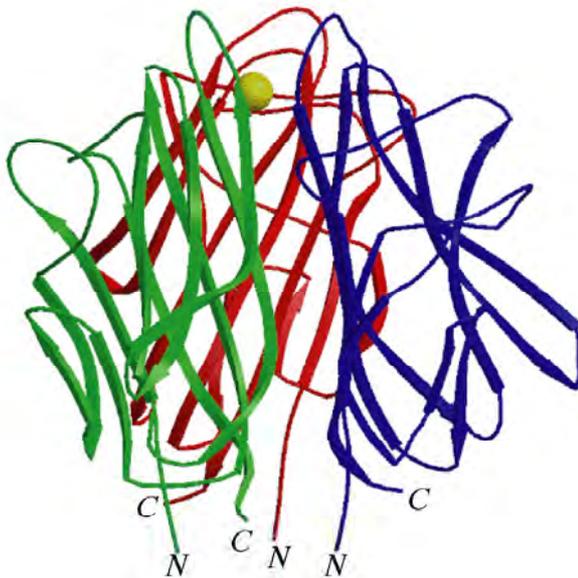


Figure C.2.1. Structure de la tête globulaire de C1q, résolue à 1.9 Å de résolution, vue de côté. On notera la forme sphérique et la présence d'un ion calcium (jaune) au sommet. Dans la molécule complète, une triple hélice de type collagène précède ce domaine en N-terminal. Chaînes A (bleu), B (vert) et C (rouge). (publi n° 14)

La structure hétérotrimérique des têtes globulaires (Figure C.2.1, publi n°14) apparaît extrêmement compacte, presque sphérique, comparée à la structure de la MBL, ouverte en forme de trèfle (cf chapitre F). L'interface entre les 3 chaînes est maintenue essentiellement par des liaisons non polaires à la base, puis une ouverture se fait entre les trois sous-unités, créant une dépression centrale dans laquelle nous avons observé la présence, inattendue, d'un atome de calcium (Figure C.2.1). Contrairement à la MBL et aux ficolines, les séquences des trois chaînes de C1q sont différentes, ce qui se traduit par des propriétés de surface variées pour chaque sous-unité (Figure C.2.2).

Il est donc vraisemblable que la variété de signaux que ce domaine est capable de détecter (cf B.2.2) résulte de leur fixation sur des sites de liaison différents, qui peuvent être localisés sur des faces variées de la molécule C1q. Comme la structure est très compacte, on peut aussi envisager l'existence d'un ou plusieurs sites d'interactions à cheval sur deux ou trois sous-unités voisines. Plusieurs molécules polyanioniques ont été identifiées parmi les premiers ligands observés pour C1q, et donc il a été suggéré que les têtes globulaires de C1q se fixent sur des motifs de charges. La figure C.2.2 montre en effet beaucoup de résidus chargés (en bleu, cyan, rouge et magenta) sur chacune des faces.

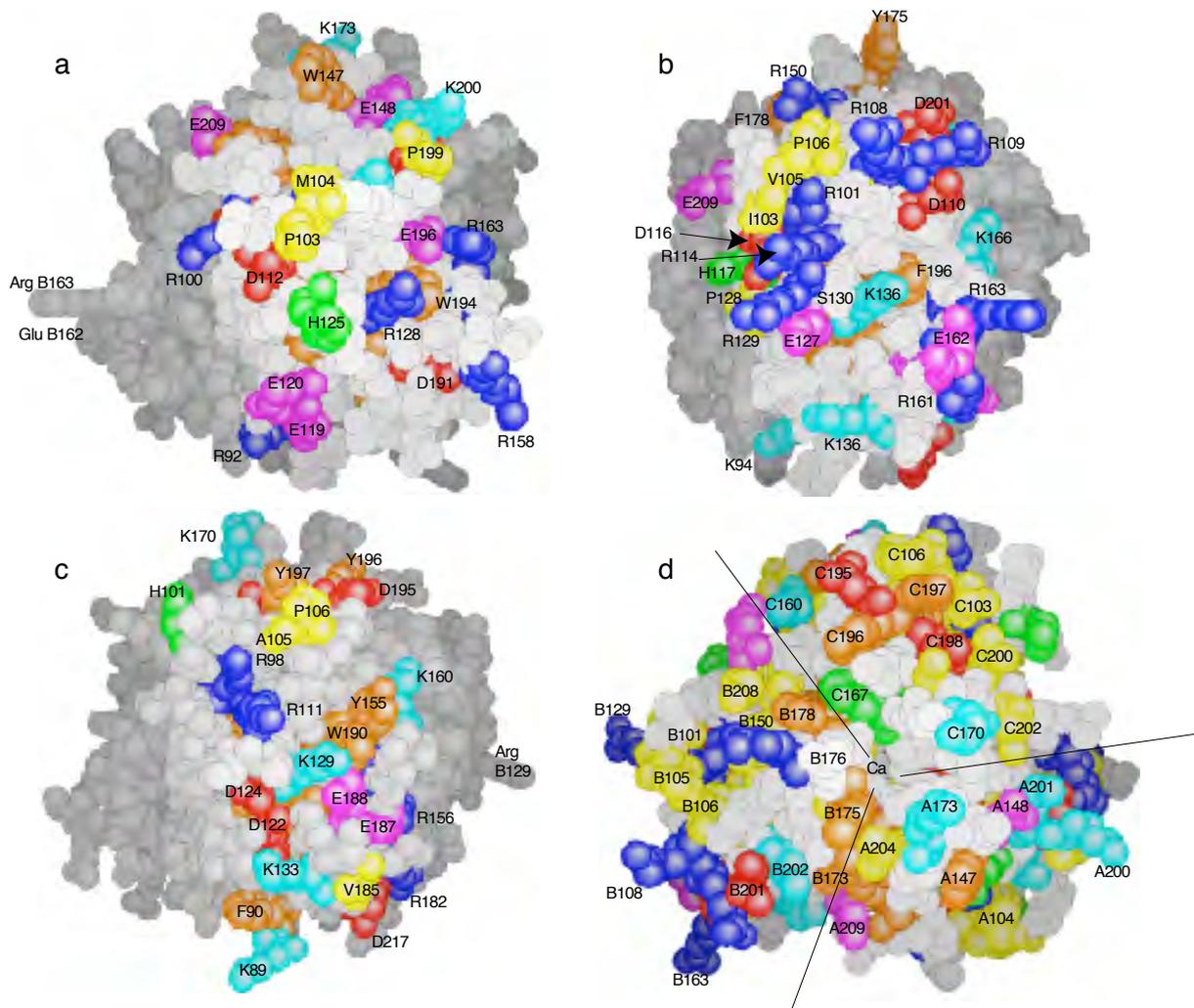


Figure C.2.2. La structure des têtes globulaires présente des faces contrastées, susceptibles d'interagir avec différents ligands: (a,b,c) vues de côté, face aux sous-unités A, B, C ; (d) vue de dessus. Les acides aminés chargés positivement sont en bleu (Arg) et cyan (Lys) (et vert pour les His) ; ceux chargés négativement en rouge (Asp) et magenta (Glu) ; les groupements aromatiques exposés sont en orange (Phe, Trp, Tyr) ; enfin les groupements de résidus hydrophobes exposés sont en jaune. Ca : calcium (*publi n° 14*)

C.2.2. Deux modèles d'interaction possibles des domaines globulaires de C1q.

Pour illustrer cette variété possible de positions des sites d'interaction, j'ai effectué un travail de modélisation pour deux exemples concrets : CRP (C Reactive Protein) et IgG1 (immunoglobine G de type 1) pour lesquels nous disposons à la fois de la structure de la protéine reconnue et d'informations sur les mutations modulant cette reconnaissance et/ou l'activation induite de la voie classique.

On peut voir ainsi, sur la figure C.2.3, un exemple d'interaction mettant en jeu les trois sous-unités simultanément (interaction avec la CRP), et, sur la figure C.2.4, un exemple d'interaction latérale médiée par la partie équatoriale de la sous-unité B (interaction avec IgG1).

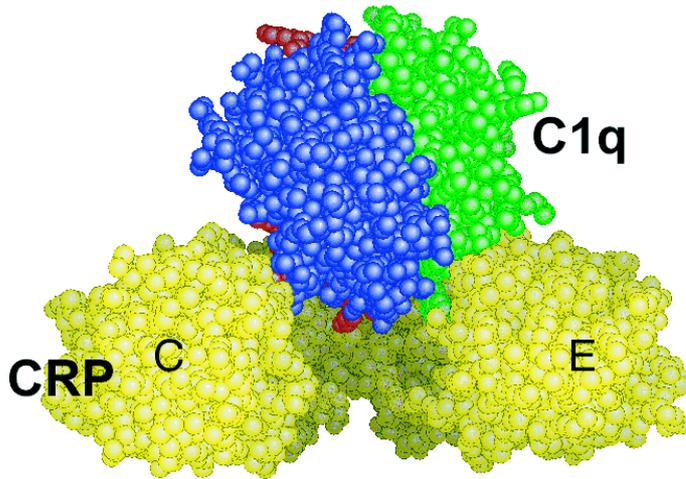


Figure C.2.3 Modèle d'interaction médiée par la face apicale de C1q avec le pentamère de la CRP (C Reactive Protein).

Deux sous-unités (A et B) de la CRP (en jaune) ont été enlevées pour montrer le canal central où les surfaces entrent en contact étroit. Les trois chaînes de C1q interagissent avec la CRP: A (bleu), B (vert) et C (rouge). (*publi n° 14*)

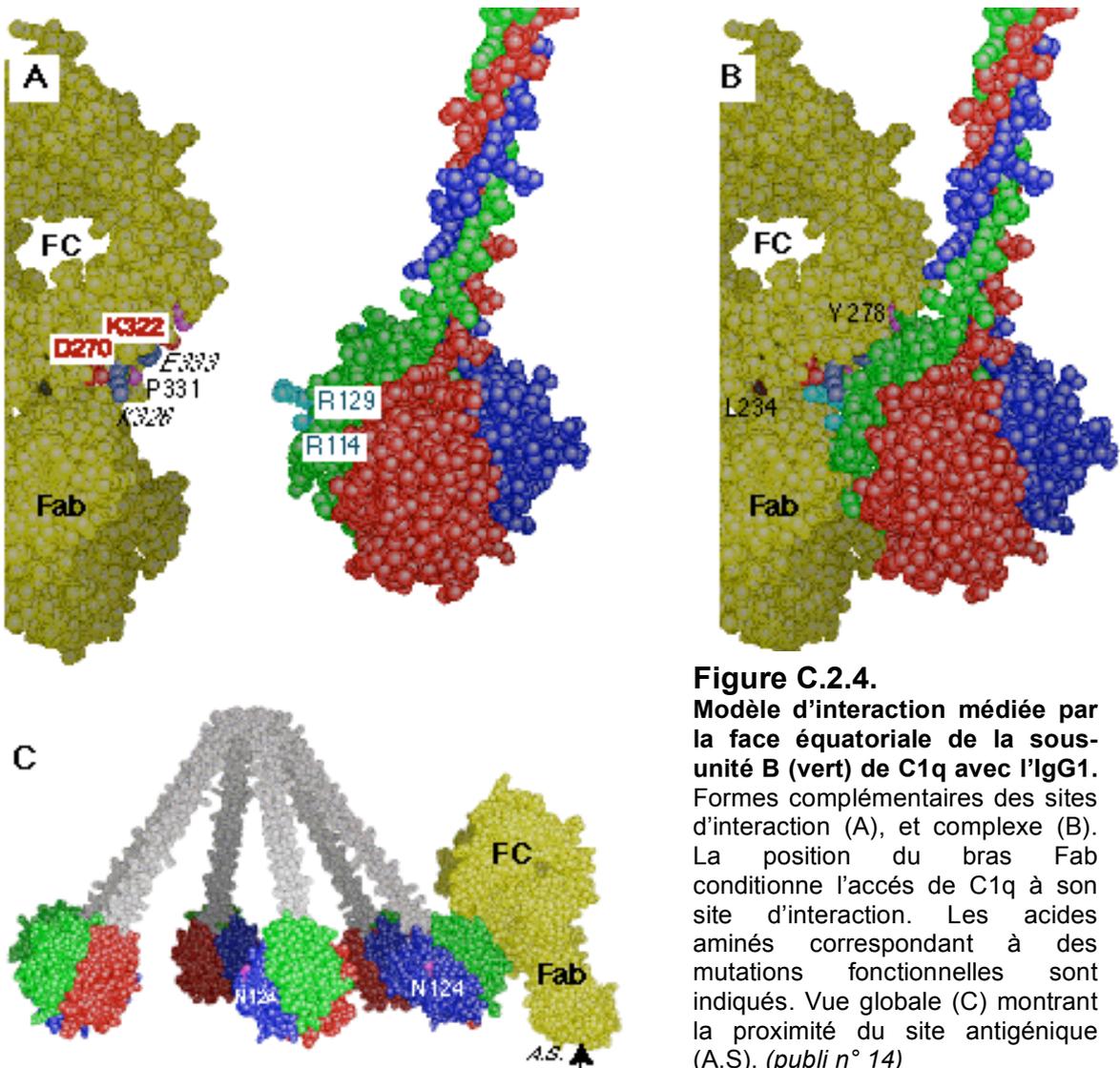


Figure C.2.4.

Modèle d'interaction médiée par la face équatoriale de la sous-unité B (vert) de C1q avec l'IgG1.

Formes complémentaires des sites d'interaction (A), et complexe (B). La position du bras Fab conditionne l'accès de C1q à son site d'interaction. Les acides aminés correspondant à des mutations fonctionnelles sont indiqués. Vue globale (C) montrant la proximité du site antigénique (A.S.). (*publi n° 14*)

Je ne reprendrai pas ici en détail l'élaboration de ces modèles, argumentée dans la publication n° 14. Je commenterai cependant le modèle d'interaction IgG1-C1q (Figure C.2.4), que j'ai élaboré en utilisant la structure de l'anticorps entier b12 (Saphire et al., 2001), en précisant qu'il est en accord avec l'ensemble des données expérimentales publiées. Ces dernières incluent une analyse biophysique de l'interaction IgG-C1q (Burton et al., 1980), la nitration des tyrosines des IgG (McCall and Easterbrook-Smith, 1989), des modifications chimiques des arginines de C1q (Marqués et al., 1993), des mutations de la charnière d'IgG1 (Guddat et al., 1993), des mutations d'anticorps IgG1 humain (Idusogie et al., 2000, 2001; Hezareh et al., 2001), et, plus récemment, des mutations de formes recombinantes chimériques des sous-unités de gC1q (Kojouharova et al., 2004). Je précise ceci pour souligner mon souci permanent d'élaborer des modèles pour les confronter aux données expérimentales: les modèles construits deviennent ainsi un excellent support pour regrouper, de façon synthétique et dans l'espace 3D, l'ensemble des informations expérimentales publiées. Cette liste de références montre également l'intérêt que suscite, depuis longtemps, l'analyse des interactions entre C1q et les immunoglobulines.

C1q est connu pour se fixer au domaine Fc des anticorps, et, plus précisément, pour des réseaux d'anticorps fixés sur une surface antigénique. Dans ce modèle, non seulement C1q contacte ce domaine Fc, mais il se rapproche aussi du bras Fab (Figure C.2.4). Ce contact supplémentaire, observé dans cette conformation particulière du bras Fab, n'a pas forcément de valeur générale ; cependant, le site de fixation des têtes globulaires de C1q semble masqué dans l'immunoglobuline libre, son accès étant libéré par le changement de position du bras Fab après fixation de l'antigène. La position relative Fc/Fab semble en effet cruciale (perte ou augmentation de la liaison de C1q pour des mutants de différentes régions de la région charnière d'IgG1 ; Dall'Acqua et al., 2006), ce qui est peut-être un moyen pour C1q de reconnaître sélectivement un anticorps fixé à un antigène.

Des projets d'ingénierie d'anticorps, développés dans l'objectif d'une thérapie anticancéreuse, ont ainsi montré la possibilité de moduler finement la fixation de C1q, par quelques mutations, et donc de moduler l'importance du système du complément comme effecteur (Idusogie et al., 2000, 2001 ; Dall'Acqua et al., 2006).

Enfin, ce modèle suggère aussi que, dans le cas d'une fixation sur une large surface (comme une bactérie), il pourrait y avoir également interaction entre cette surface et d'autres sous-unités de C1q (Figure C.6-C). En effet, les différentes têtes globulaires sont orientées en direction de la surface antigénique, qui peut aussi présenter d'autres motifs moléculaires reconnus par C1q. Ainsi, C1q est une molécule qui présente la propriété singulière de pouvoir interagir avec avidité, grâce à ses multiples têtes globulaires, avec des motifs moléculaires différents sur une même surface. Nous pensons que cette propriété de C1q est sans doute essentielle aussi pour lui permettre de se fixer directement avec avidité sur des surfaces du soi altéré (publi 34).

C.2.3. Modélisation des tiges de type collagène de C1q.

La structure des têtes globulaires m'a également fourni les informations nécessaires pour construire un modèle des tiges de type collagène au niveau des bras de C1q. En effet, sans cette structure, il était impossible de prédire sans erreur la position relative, dans le trimère, des trois chaînes distinctes A, B et C, ainsi que l'alignement correct des triplets de collagène. La première tentative de modélisation des régions de type collagène de C1q (Kilchherr et al., 1985) avait ainsi produit un modèle incompatible avec notre structure cristallographique.

Pour modéliser la triple hélice de collagène, j'ai utilisé les paramètres dérivés de l'analyse statistique de Rainey and Goh (2002), qui différencie les triplets riches ou pauvres en

proline. Cette modélisation des bras de la molécule C1q m'a permis de proposer une version plus complète du modèle de la molécule entière, ce qui servira de support pour étudier la question de la transmission du signal d'activation dans le complexe C1 (voir ci-dessous).

C.3. Domaine d'interaction de la protéase C1s.

Les protéases C1r et C1s s'associent sous la forme d'un tétramère C1s-C1r-C1r-C1s, et se fixent à la molécule C1q pour former le complexe C1. Ces interactions font intervenir les extrémités N-terminales (CUB1-EGF) des deux protéases, qui interagissent entre elles sous la forme d'un hétérodimère. Il paraissait important d'essayer de déchiffrer les bases structurales de ces interactions, pour pouvoir mieux en apprécier les propriétés 'mécaniques'. Nous avons ainsi essayé de savoir comment ces interactions ternaires étaient possibles. Une question importante était aussi de savoir sous quelle forme le signal de fixation des têtes globulaires de C1q sur différentes cibles un peu éloignées pouvait se transmettre aux protéases C1r et C1s associées ? Nous avons exploré ces questions à partir de l'étude structurale du domaine d'interaction de C1s, le fragment homologue de C1r étant difficile à produire en quantité suffisante pour une étude structurale.

C.3.1. Structure du fragment CUB1-EGF de C1s.

La structure des modules CUB1-EGF de C1s a été résolue et affinée à 1.5 Å de résolution (publi n° 15). Cette structure a été obtenue sous forme dimérique en présence de calcium ou magnésium.

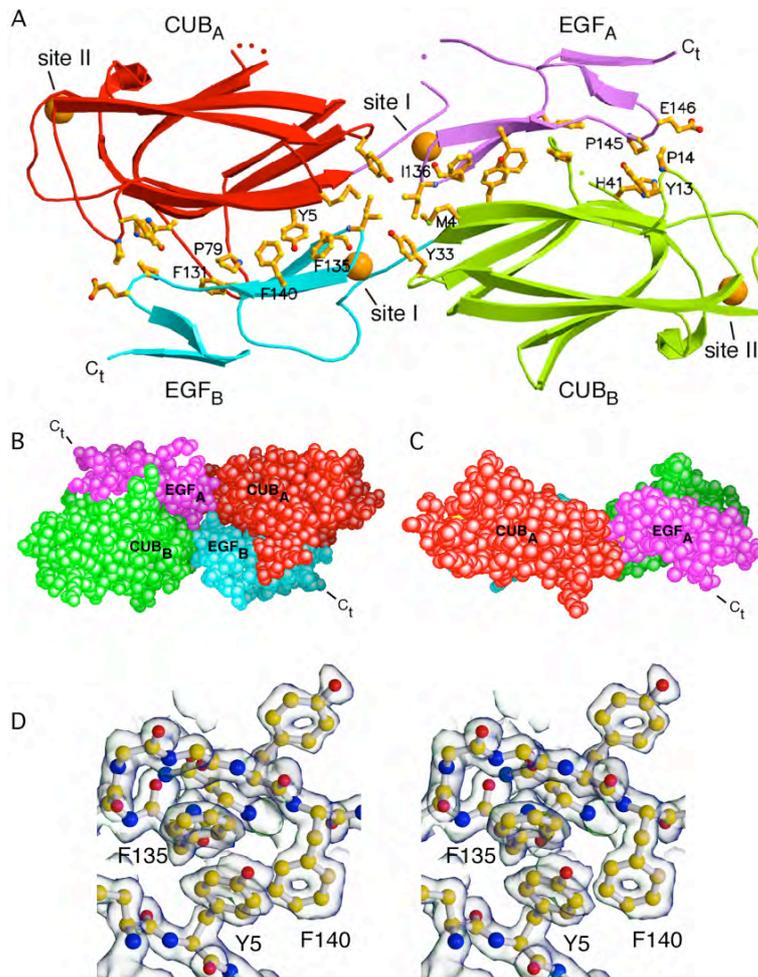


Figure C.3.1.
CUB1-EGF : domaine d'interaction de C1s, dépendant du Ca²⁺

A) Vue générale, de dessus, avec le détail de l'interface stabilisant le dimère sous forme tête-bêche, et les positions des sites de fixation du calcium (I et II, par monomère). B) Vue de dessus, en sphères pleines, montrant les contacts étendus et étroits entre les modules CUB et EGF de chaque monomère. C) Vue de côté, montrant un sillon sur le dessus, dont la taille correspond au diamètre d'une triple hélice de type collagène. D) Vue détaillée de la densité électronique à 1.5 Å de résolution, au niveau de l'interface du dimère (trépiéd aromatique). Carte 2mFo-DFc contournée à 1 σ en vue stéréoscopique (publi n° 15).

Pourquoi la présence du calcium est-elle essentielle pour la formation du dimère ? Les deux molécules s'assemblent de façon tête-bêche (Figure C.3.1), le module CUB d'un monomère s'associant au module EGF de l'autre monomère. Cette interaction est fortement stabilisée par les ions calcium fixés au centre du dimère au niveau de chaque module EGF. Ces ions calcium stabilisent à la fois l'interface entre les modules CUB et EGF du même monomère et l'interface du dimère, entre un module CUB et le module EGF de son partenaire.

C.3.2. Modèle d'association des domaines d'interaction de C1s et C1r avec C1q.

Cette structure nous a permis de proposer un modèle pour l'hétérodimère physiologique (CUB1-EGF de C1r et CUB1-EGF de C1s), également dépendant du calcium. En effet, les acides aminés stabilisant l'interface de l'homodimère CUB1-EGF de C1s sont très fortement conservés dans les séquences de C1r et des protéases MASPs. D'ailleurs une association dimérique similaire a été observée dans la structure de domaine N-terminal CUB1-EGF-CUB2 de la protéine MASP2 de rat (Feinberg et al., 2003), résolue à la même époque. Nous avons pu identifier dans la structure de C1s un site inattendu de fixation du calcium dans chaque module CUB1 (site II, figure C.3.1). Ce site a au moins un rôle de stabilisation structurale dans C1s. Ce site II sera retrouvé dans les protéases MASPs, dans lesquelles nous avons pu démontrer un rôle fonctionnel. Il sera donc présenté dans ce contexte (paragraphe D.1.1)

Comment associer le modèle de l'hétérodimère C1r-C1s avec les tiges de type collagène de C1q (modélisées après résolution de la structure des têtes globulaires de C1q) ? Sur la vue latérale (C) de la figure C.3.1., on peut remarquer la présence d'un sillon dont la taille correspond environ au diamètre d'une triple hélice de collagène. C'est sur cette hypothèse qu'un modèle d'association C1r/C1s/C1q a été construit. Là encore, nous avons pris en compte les résultats précédents de caractérisation biochimique des interactions entre C1q, C1r et C1s obtenus après diverses modifications chimiques (Illy et al., 1993). Nous avons également analysé minutieusement la disposition des modifications post-traductionnelles observées sur les tiges de C1q (Reid, 1979) et les particularités de séquence du module EGF de C1r. Le modèle obtenu est montré sur la Figure C.3.2. Un retour sur les Figures C.1.1 et C.1.2 aidera à situer ce fragment dans le complexe C1 entier.

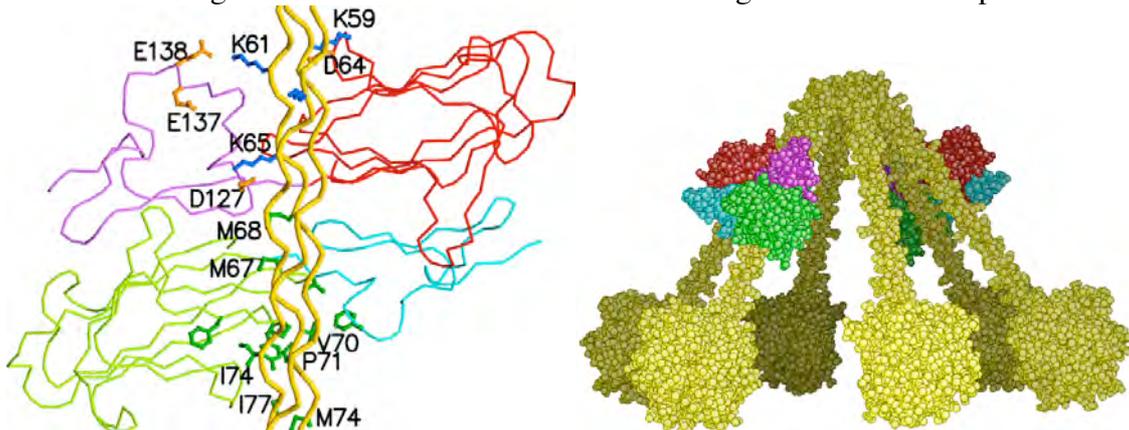


Figure C.3.2. Hypothèse d'interaction entre les bras de C1q (jaune) et les parties N-terminales CUB-EGF de l'hétérodimère C1r (rouge-magenta) - C1s (vert-bleu). A gauche, le modèle est détaillé autour des sites probables d'interaction, comprenant des résidus chargés (en orange) du côté de C1r, et des acides aminés hydrophobes du côté de C1s. Les lysines A59 et B61 de C1q (en bleu) ne sont pas modifiées, la lysine B65 est hydroxylée mais non glycosylée. A droite, ces interactions sont placées dans le contexte du cône de C1q, les protéases s'enroulant autour des tiges de C1q pour se fixer sur leur face extérieure (*publi n° 15*).

C.4. Domaine catalytique de la protéase C1r.

La fixation du complexe C1q sur une surface activatrice peut déclencher, dans des conditions induisant son étirement, une cascade d'activations internes. La séquence d'évènements identifiée est la suivante : d'abord les protéases C1r sont activées, par coupure d'une liaison spécifique de leur site d'activation, puis elles activent ensuite la protéase C1s, également par coupure d'une liaison spécifique de son site d'activation. Cependant, la question de savoir comment une telle cascade d'activation peut se produire à l'intérieur du complexe C1 restait énigmatique.

C.4.1. Structure de l'assemblage dimérique des domaines catalytiques de C1r.

La structure de l'assemblage dimérique des domaines catalytiques de C1r a été résolue à 2.9 Å de résolution (publi n° 17). Ce dimère de domaines catalytiques de C1r correspond à la partie centrale du tétramère C1s-C1r-C1r-C1s, logée au centre du cône de C1q (Figures C.1.1 et C.1.2). C'est ce domaine catalytique qui permet à la fois l'activation de C1r, puis celle de C1s.

La structure affinée du fragment dimérique CCP1-CCP2-SP de C1r a été obtenue sous forme proenzyme, stabilisée par une mutation (Arg446Gln) du site d'activation. La structure proenzyme est globalement de même type que celle du chymotrypsinogène, mais présente des particularités locales significatives, en particulier la présence d'une hélice de surface qui n'existe que dans C1r, et qui contribue à la stabilisation du dimère.

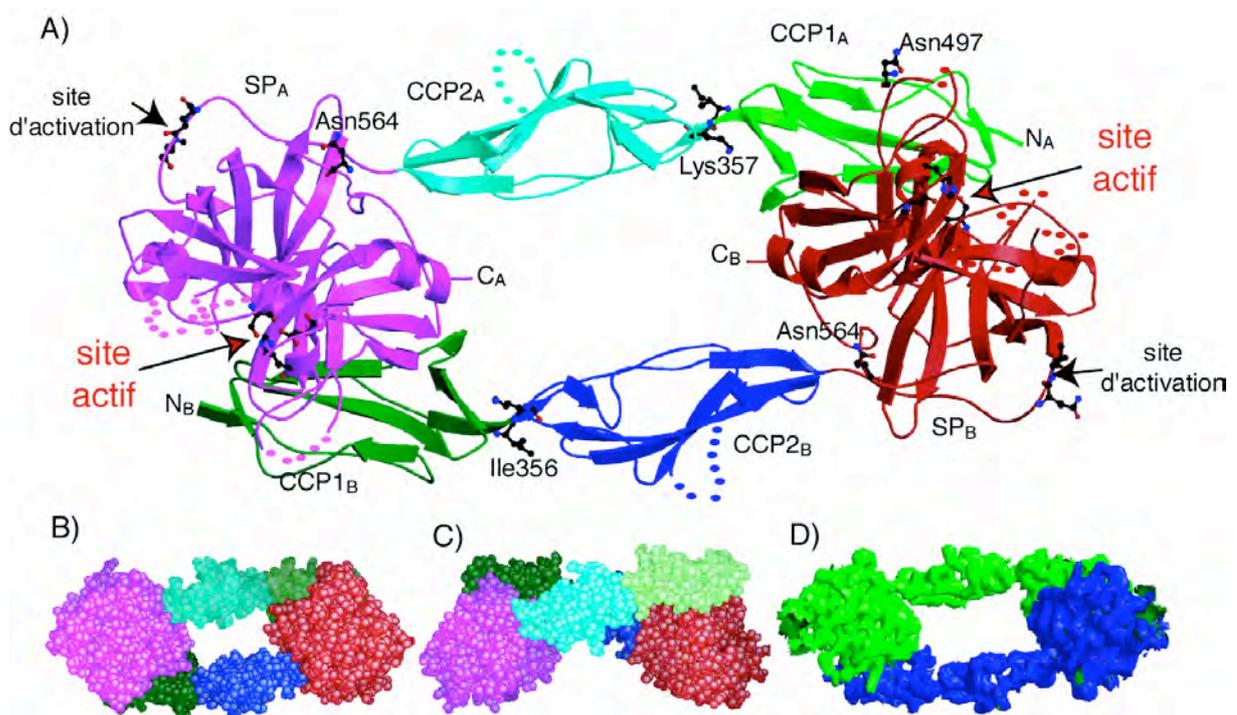


Figure C.4.1. Structure dimérique du domaine catalytique CCP1-CCP2-SP de C1r . A) Vue en ruban, montrant la structure tête-bêche, stabilisée par des contacts entre le module CCP1 d'un monomère avec le domaine SP de l'autre monomère. Dans cette structure « au repos », le site actif (a.s.) d'un monomère est très éloigné du site d'activation de l'autre monomère : aucune activation mutuelle n'est possible dans cette configuration! B) Vue en sphères pleines (même orientation qu'en A). On remarque une grande ouverture centrale. C) Vue de côté. D) Densité à moyenne résolution (4Å) observée pour un domaine catalytique activé de C1r : même structure de dimère « au repos ». (publi n° 17)

La structure dimérique présente une forme allongée (116 Å x 56 Å), les deux molécules étant disposées de façon ‘tête bêche’ (Figure C.4.1) : on observe une interaction non covalente quasi-symétrique entre le premier module CCP1 d’un monomère et le domaine SP de son partenaire. Cette structure dimérique est cohérente avec l’ensemble des données de biochimie, de biologie moléculaire et de microscopie électronique préexistantes, donc très vraisemblablement physiologique. Par exemple, l’expression préalable de fragments de différentes longueurs de ce domaine a montré que seuls les fragments comprenant les trois modules CCP1, CCP2 et SP étaient obtenus sous forme de dimères, des monomères étant obtenus en absence du module CCP1 (Lacroix et al., 2001). Une association similaire a été observée à moyenne résolution (4 Å) sur un fragment CCP1-CCP2-SP de C1r sous forme activée (Figure C.4.1.D).

C.4.2. Implication pour le mécanisme d’activation : un étirement est nécessaire.

L’élément le plus surprenant a été de trouver les sites actifs et sites d’activation de chaque monomère aux deux extrémités du dimère, donc très distants (>90Å) l’un de l’autre... Cette configuration est incompatible avec l’étape d’auto-activation initiale de la protéase C1r. En effet, il faut dissocier le dimère observé ici pour mettre les deux protéases dans une orientation relative compatible avec l’activation d’une molécule (avec le site d’activation d’une molécule introduit dans le site actif de l’autre molécule, voir l’étape n°3 de la figure C.4.2). Néanmoins, la structure du dimère montre une large ouverture centrale (30 Å x 13 Å, voir figure C.4.1), ce qui n’avait pas pu être envisagé auparavant. Cette ouverture va permettre le glissement des domaines catalytiques de C1r et leur réorientation. Ceci permet d’envisager un nouveau processus d’activation de C1r dans le complexe C1, dans lequel les deux domaines catalytiques peuvent glisser l’un sur l’autre (Figure B.3.10, étapes 1 à 3). Ce mouvement pourrait être induit par un stress mécanique transmis par les extrémités N-terminales des protéases, ancrées sur la face extérieure des tiges de C1q, lors d’une déformation du cône de C1q. En effet, une telle tension mécanique provenant du mouvement des têtes globulaires de C1q, transmise jusqu’aux extrémités N-terminales de chaque domaine catalytique, semble correctement orientée pour pouvoir provoquer la rupture du dimère et favoriser une activation de C1r (Cf Figures C.4.2 et C.6.1). Après activation de C1r et C1s, le dimère activé peut reprendre sa conformation stable au repos (Figure C.4.1.D).

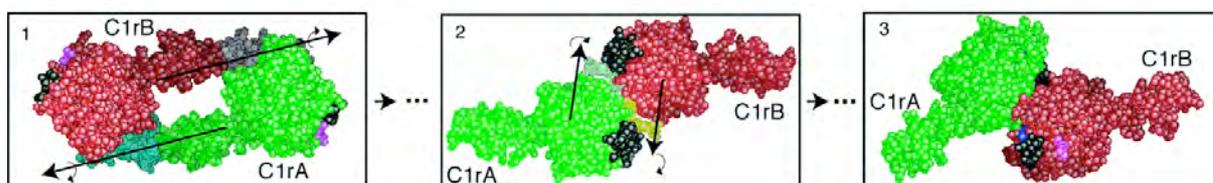


Figure C.4.2. L’activation de C1r, dans le contexte du complexe C1, nécessite la rupture du dimère par étirement de ses extrémités.

- 1) Dimère au repos. Les flèches symbolisent la traction mécanique que peuvent exercer les extrémités N-terminales des protéases en cas de déformation du cône de C1q, ce qui conduira aux étapes suivantes. 2) Contact symétrique entre deux molécules de C1r zymogène forme courte : ceci pourrait représenter un état intermédiaire sur le chemin d’activation ; 3) Association de type Enzyme-Produit vue dans le cristal de la forme courte activée de C1r : C1rB est en position d’activation de C1rA (*publi n° 16*)

C.4.3. Structures des formes courtes CCP2-SP du domaine catalytique de C1r.

Grâce à une collaboration industrielle avec BASF, nous avons pu obtenir également la structure de la forme active de C1r, à 2.8 Å de résolution, sur un fragment plus court (CCP2-SP). Ceci nous a permis d’étudier les bases structurales de la spécificité de C1r,

d'observer précisément les changements conformationnels entre forme zymogène et forme activée (publi n° 16). Enfin, en regardant les empilements cristallins des fragments courts CCP2-SP, monomériques, nous avons pu observer une interaction de type 'enzyme-produit' pour la forme sauvage active, et une interaction symétrique 'face à face' de deux zymogènes pour le mutant Ser->Ala du site actif. Ces deux types d'interaction pourraient bien mimer des interactions intermédiaires lors du processus d'activation de C1r (figure C.4.2).

C.5. Structure du domaine catalytique de C1s

La première structure résolue a été celle du fragment CCP2-SP du domaine catalytique de C1s (publi n°18), responsable de la coupure et activation des substrats spécifiques C4 et C2 du complément humain. Il est important de noter que le passage de la purification de fragments produits par protéolyse ménagée à la production de fragments recombinants produits dans le système de type baculovirus a été un élément important de ce premier succès.

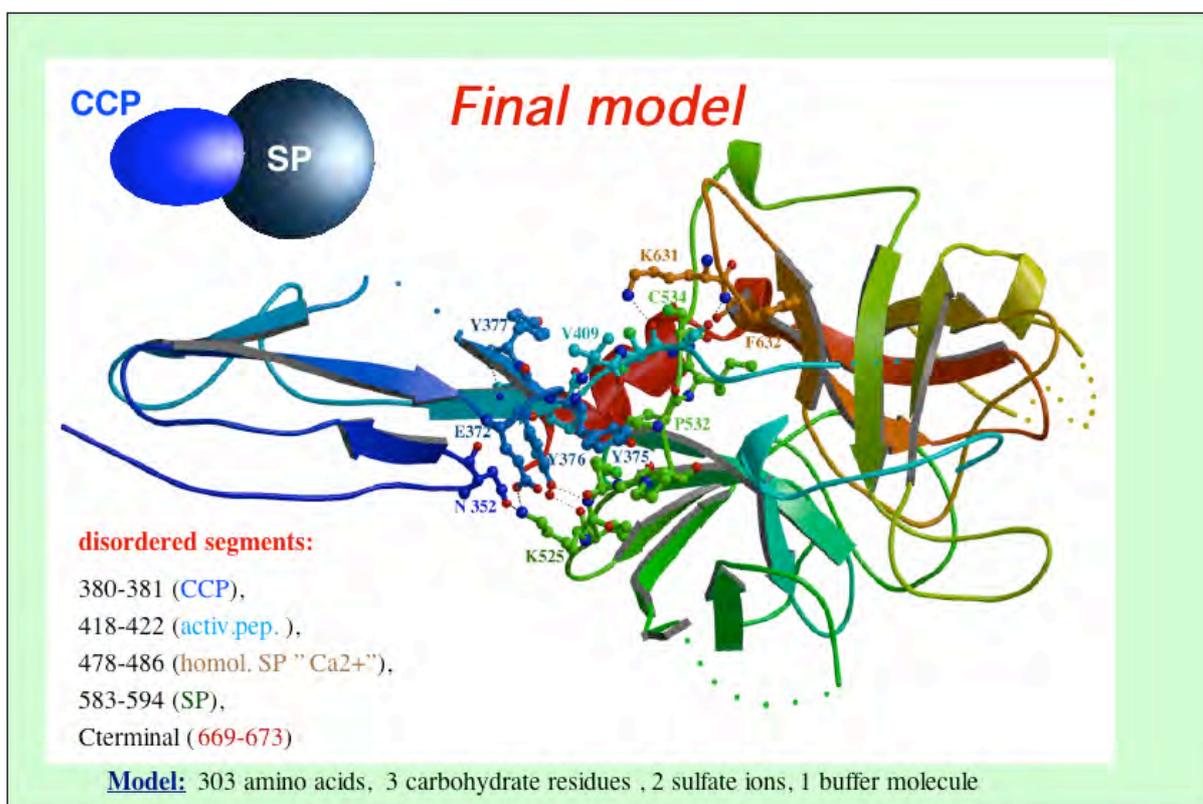


Figure C.5.1. Structure globale du domaine catalytique CCP-SP de C1s : une protéase à manche. Cette figure montre l'orientation perpendiculaire du module CCP à la surface du domaine SP (ellipsoïdes) ainsi que les acides aminés stabilisant cet ancrage à l'interface module-domaine. Les quelques régions désordonnées de cette structure sont aussi définies et localisées. (adapté d'après la publi n° 18)

Cette structure a été résolue et affinée à 1.7 Å de résolution. Nous avons constaté que le fragment catalytique CCP2-SP avait cristallisé sous forme activée, même s'il avait été produit sous forme recombinante purement zymogène. Les trois premiers résidus du site de glycosylation sur le module CCP2 étaient clairement visibles.

C.5.1. Déterminants structuraux de la spécificité de la protéase C1s.

La spécificité très étroite de C1s pour ses deux substrats, C4 et C2, et pour son inhibiteur physiologique (C1-inhibiteur) est déterminée en partie par un ensemble de sites secondaires de liaison du substrat dans le domaine SP. La structure a permis de déterminer la structure précise de ces sous-sites (Figure C.5.2). De plus, il apparaît clairement que l'accès à la plupart de ces sous-sites est extrêmement restreint par des contraintes stériques (Figure C.5.2).

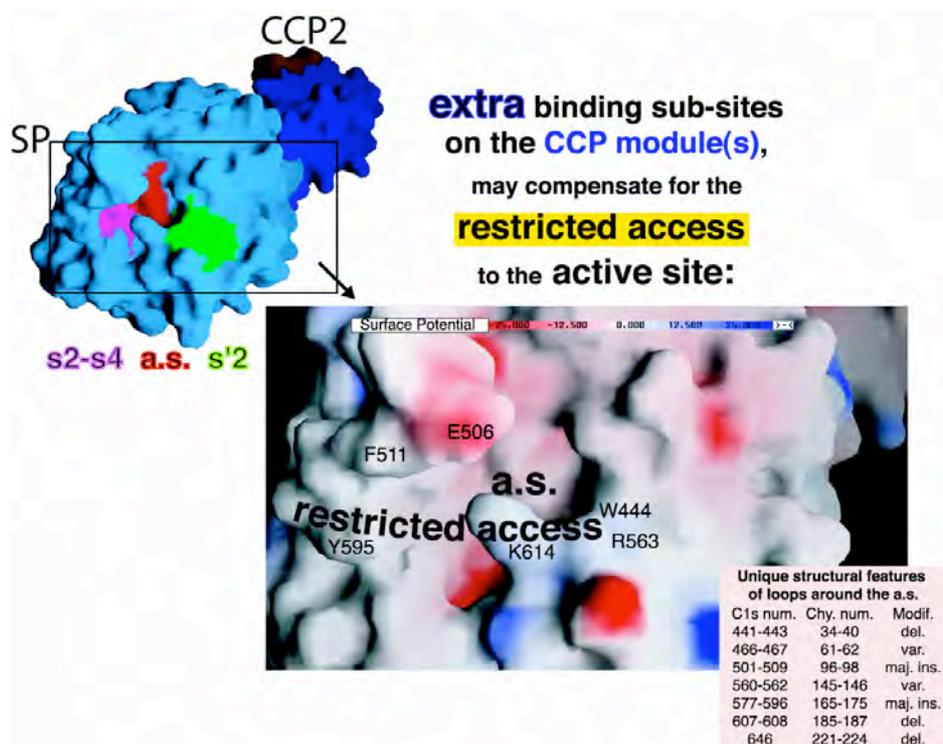


Figure C.5.2. Accès restreint au site actif de C1s : nécessité de sites de liaison additionnels fournis par les modules CCP (en particulier CCP1).

La vue en gros plan du site actif (a.s.) montre son accès restreint, la lysine 614 masquant en particulier l'accès du site principal. La vue globale (surfaces bleues colorées) montre l'emplacement du site actif (rouge) et des sous-sites de liaison des acides aminés présents de part et d'autre de la liaison peptidique à cliver (violet et vert). Le module CCP2 (et certainement également le module CCP1) offrent un prolongement en arrière plan à ce site de liaison du domaine protéase à sérine (SP). Les boucles de surface définissant la spécificité fine de C1s sont indiquées en bas à droite. (*adapté d'après la publi n° 18*)

Les tests d'activité réalisés sur différents fragments recombinants de C1s (SP, CCP2-SP, CCP1-CCP2-SP ; Rossi et al., 1998), ainsi que sur des mutants de la charnière CCP1-CCP2 (publi n°11), démontrent clairement la présence d'éléments de reconnaissance additionnels sur le module CCP1, nécessaires au positionnement correct du substrat C4 pour son clivage. L'orientation particulière du module CCP2 par rapport au domaine SP est tout à fait compatible avec cette notion d'extension du site de liaison du substrat C4 par le module CCP1, et l'accès limité aux sous-sites du domaine SP devrait être facilité par cette fixation additionnelle sur le module CCP1.

C.5.2. L'orientation perpendiculaire du module CCP2 à la surface de C1s: une fonction de « levier » à l'intérieur du complexe C1 ?

Une autre surprise de cette structure a été l'orientation du module CCP2, ancré perpendiculairement à la surface de C1s (Figure C.5.1). Une rigidité de l'interface entre le domaine SP et le module CCP2 avait été prédite à partir d'une analyse originale des

séquences (voir A.2.3), mais sans toutefois présumer d'une orientation relative perpendiculaire du domaine et du module. Cette analyse de séquences (publi n° 19) prédisait une conservation de cet assemblage dans une sous-famille de protéines, ce qui a été confirmé *a posteriori* dans le cas de l'assemblage entre le module CCP2 et le domaine SP de C1r et de MASP2.

Cette orientation très particulière du module CCP2 peut avoir des implications fonctionnelles dans le contexte du complexe C1. Le domaine SP de C1s doit en effet passer d'une position initiale interne, nécessaire à son activation par C1r, à une position plus externe, lui donnant accès à ses substrats protéiques C4 et C2 (Figure C.5.3). Il est probable que ce mouvement soit rendu possible par l'existence de charnières flexibles, par exemple entre les modules CUB2 et CCP1, et/ou CCP1/CCP2. De par son orientation perpendiculaire au domaine SP, le module CCP2 agit comme un bras de levier et amplifie le déplacement du domaine SP (Figures C.5.3 et C.5.4). La structure cristallographique du domaine catalytique de C1s, la première décrite pour une protéase modulaire du système du complément, a donc apporté des éléments imprévus et précis sur les déterminants structuraux de la spécificité de C1s et de son activité dans son contexte physiologique.

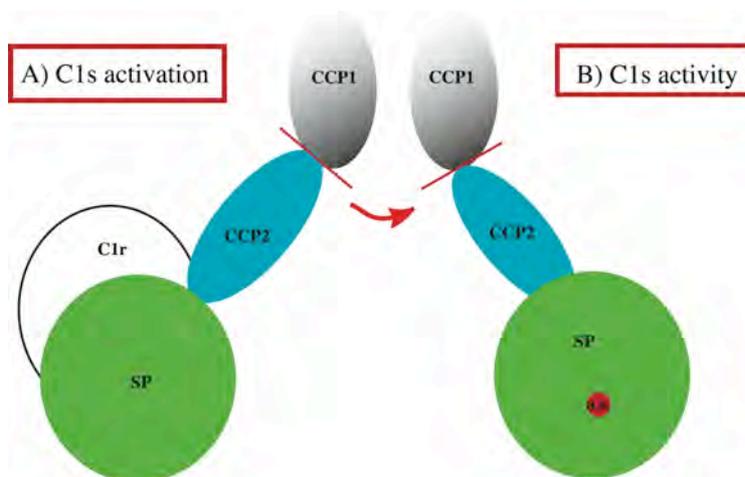


Figure C.5.3. Un rôle possible de bras de levier pour le module CCP2. Des déplacements et/ou torsions assez fins en N-terminal du domaine catalytique (ou comme ici, entre les modules CCP1 et CCP2) se trouvent amplifiés par la présence de ce « manche », ce qui permet au domaine catalytique de C1s de passer d'une localisation interne au complexe C1 (dessin de gauche), nécessaire à l'étape de son activation, à une position plus externe (dessin de droite), le site actif étant orienté vers l'extérieur (rond rouge marqué a.s.). C'est dans cette deuxième position qu'il pourra cliver les substrats C4 puis C2. C1r marque une position interne au complexe C1, alors que CCP1 marque la limite externe, la charnière CCP1-CCP2 pouvant s'appuyer sur une tige collagène de C1q (*adapté d'après la publi n° 18*)

C.5.3. Exploitation de la structure pour le développement d'inhibiteurs.

La résolution de cette structure nous a permis également d'établir une collaboration avec BASF, cette entreprise étant intéressée par la conception d'inhibiteurs spécifiques de C1s. Des structures de C1s en complexe avec des inhibiteurs ont été réalisées par nos collaborateurs, mais elles n'ont jamais été publiées. D'autres chercheurs se sont également intéressés au développement d'inhibiteurs spécifiques de C1s. Deux séries de développements ont été publiées par Nalin Subasinghe. Ces développements ont été basés sur l'analyse de la structure de

C1s et la modélisation de complexes formés avec des petits inhibiteurs naturels et leurs dérivés. De nombreux dérivés ont été synthétisés et analysés, ces molécules différant par la nature chimique de groupements latéraux sur un squelette de base commun : les dérivés « thiopheneamidines » (Subasinghe et al., 2004) et les dérivés du motif « arylsulfonylthiophene-2-carboamidine » (Subasinghe et al., 2006). Une des difficultés est d'obtenir un inhibiteur vraiment spécifique de C1s, les protéases à sérine de structure voisine étant nombreuses dans le plasma: thrombine, facteur D, activateur plasminogène tissulaire, urokinase plasminogène activateur (uPA), plasmine, MASP2, facteur B, C2, La présence atypique de la Lysine 614 à l'entrée du site actif de C1s (Figure C.5.2) a été exploitée avec succès dans la deuxième étude (Subasing et al., 2006) pour améliorer la sélectivité des inhibiteurs pour C1s, en particulier par rapport à l'uPA. Réciproquement, l'inhibiteur BCX-1470, développé sur la base de la structure du facteur D complexé au DFP, s'est avéré finalement un des meilleurs inhibiteurs de C1s ($IC_{50} = 1.6 \text{ nM}$) et traverse maintenant les phases d'essais cliniques (Bureeva et al., 2005).

C.6. Conclusion sur la voie classique.

Nous avons réalisé un travail pionnier sur l'étude structurale du complexe C1, qui s'est avéré un sujet de recherche passionnant sous bien des aspects. Même si d'autres tentatives d'analyses structurales de fragments de C1 ont été entreprises dans d'autres laboratoires, nous avons réussi jusqu'à présent à devancer nos concurrents.

La résolution de la structure du domaine catalytique de C1s nous a permis d'établir une collaboration avec des industriels. Cette collaboration, motivée par la conception d'inhibiteurs spécifiques de C1s, nous a aidé à progresser de notre côté sur l'analyse structurale de la protéase C1r, en renforçant les moyens financiers et humains.

L'ensemble de ces études structurales, complétées par des données de biologie moléculaire, nous a permis de proposer un modèle fonctionnel beaucoup plus précis du complexe C1, qui possède une cohérence globale avec les informations expérimentales.

Sur la base de ce modèle, nous pouvons proposer un mécanisme de transmission du signal d'activation, depuis la reconnaissance initiale de motifs moléculaires par plusieurs têtes globulaires de C1q, jusqu'à l'activation finale de la protéase C1s et des substrats C4 et C2 (cf Figure C.6.1). Selon cette hypothèse, le complexe C1 présente la capacité tout à fait remarquable de convertir un stress mécanique, généré par la reconnaissance de motifs moléculaires distants, en une cascade de réactions enzymatiques.

Les principales incertitudes qui subsistent concernent les conformations des charnières entre certains modules et leur degré de flexibilité, ainsi que la flexibilité des bras de C1q. Pour poursuivre cette approche, nous pouvons avoir recours à la Résonance Magnétique Nucléaire (RMN), pour obtenir des informations structurales sur des paires de modules, ou à la microscopie (électronique ou à force atomique), pour obtenir des enveloppes expérimentales du complexe C1, à moyenne résolution, dans lesquelles il serait possible de définir la position des structures cristallographiques des différents fragments connus. Néanmoins la flexibilité du complexe C1 et son manque de symétrie (laquelle est exagérée sur les figures) rendent le traitement des images de microscopie plus compliqué que pour des virus pour lesquels la symétrie est bien définie.

Le modèle proposé permet également de mieux percevoir le rôle crucial de l'étape de reconnaissance par les têtes globulaires de C1q dans le déclenchement du système. La nécessité d'un stress mécanique semble pouvoir limiter la probabilité d'une activation non contrôlée, ce qui est certainement renforcé '*in vivo*' par la présence de C1 inhibiteur. En effet, ce modèle implique une reconnaissance simultanée de motifs par plusieurs têtes globulaires, en accord avec la propriété de reconnaissance par avidité collective. De plus,

une tension mécanique ne peut être obtenue que si les motifs reconnus sont solidaires d'une même surface et suffisamment éloignés. Il y a donc des contraintes importantes quant à la disposition en surface des motifs reconnus, ce qui constitue un niveau supplémentaire de différenciation entre des structures étrangères, du soi 'abîmé' et du soi 'normal'.

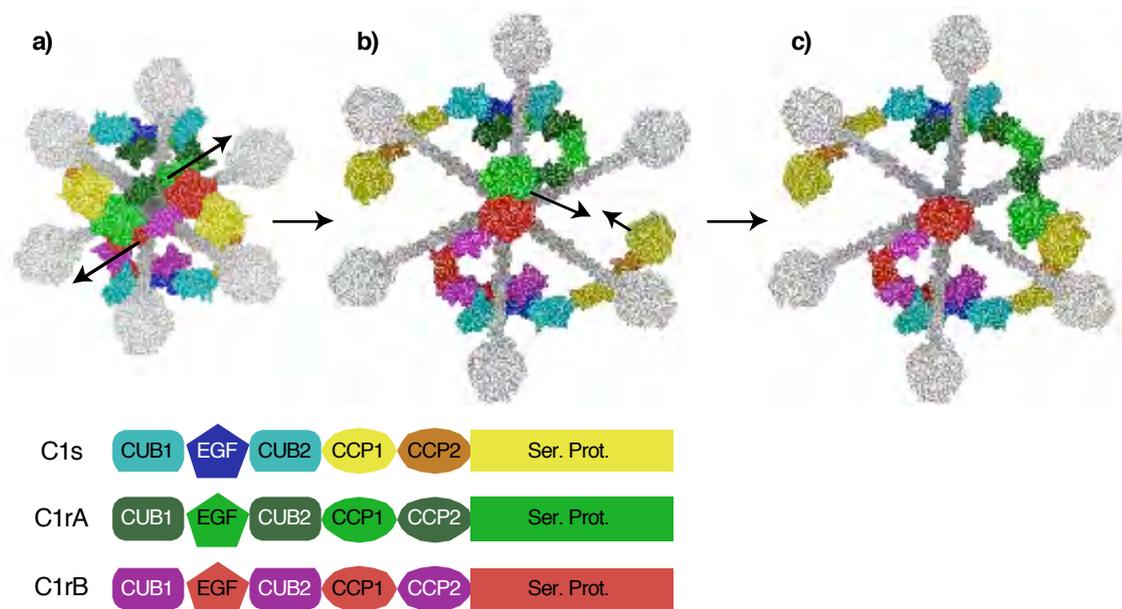


Figure C.6.1. Mécanisme hypothétique de conversion d'un signal de reconnaissance multiple en une cascade d'activations protéolytiques.

En a), le complexe C1 est au repos, vu de dessous. La reconnaissance de signaux par plusieurs têtes globulaires de C1q va entraîner une déformation avec étirement du cône de C1q sur la surface activatrice. En b), l'étirement a provoqué la dissociation du dimère de C1r central, permettant le contact entre les domaines protéases à sérine des deux molécules de C1r et leur activation. En c), il y a rapprochement entre les domaines catalytiques de C1rA et de C1s, conduisant à l'activation d'une molécule de C1s (le même phénomène pouvant se produire du côté de C1rB). La molécule de C1s activée pourra ensuite basculer vers l'extérieur pour couper le substrat C4 ; la position externe des modules CCP1 et CCP2 de C1s est compatible avec leur implication dans la reconnaissance de C4. L'ensemble du processus est caractérisé par une évolution des contacts établis par les domaines protéases à sérine de C1r et C1s. (*adapté d'après la publi n° 2*)

Les projets en cours sur cette thématique visent à tester ou améliorer ce modèle d'activation et mieux comprendre des interactions moléculaires mises en jeu dans l'étape initiale de reconnaissance, qui permet l'activation ciblée du système du complément.

D. Etudes structurales des complexes d'initiation de la voie lectine du complément.

Les dernières études structurales effectuées, concernant la MAp19 (publi n°13), l'assemblage CUB-EGF-CUB de MASP1/3 et le domaine globulaire des ficolines (publi n° 10), ont contribué à améliorer la connaissance des complexes activant cette fois le système du complément par la voie lectine. Deux types de questions ont été abordées au cours de ces dernières études :

◇ Quels sont les éléments structuraux permettant l'assemblage des complexes MBL/MASPs ou ficolines/MASPs ? Pour cela, les études cristallographiques des domaines d'interaction de MASP2 (MAp19) et MASP1/3 ont été combinées à des expériences de mutagenèse dirigée permettant de vérifier et de préciser l'identification des sites d'interaction avec la MBL et/ou les ficolines.

◇ Quels sont les éléments structuraux permettant la reconnaissance de « signaux de danger » par ces protéines ? Dans le cas des ficolines, la nature de ces signaux de danger consiste essentiellement en des motifs polysaccharidiques particuliers, se trouvant plus particulièrement à la surface de certains micro-organismes ou à la surface des cellules apoptotiques. Nous avons donc procédé à de nombreuses études cristallographiques des domaines globulaires des ficolines, cette fois-ci en essayant de visualiser leur interaction avec des ligands ou des avec analogues simplifiés de ligand.

D.1. Domaines d'interaction des protéases MASPs.

Comme pour les protéases C1r et C1s, le domaine d'interaction des MASPs est situé dans la partie N-terminale. Il conditionne la formation d'homodimères et leur association avec les tiges de type collagène de la MBL ou des ficolines. Même si les deux premiers modules CUB1-EGF peuvent permettre cette association (par exemple dans le cas de la protéine MAp19), l'affinité avec la MBL ou les ficolines est plus forte en présence du dernier module CUB2 (Cseh et al., 2002).

D.1.1. Structure de la protéine MAp19.

La protéine MAp19 est un produit d'épissage alternatif du gène *MASP2/MAP19*, comprenant les modules CUB1-EGF de MASP2, auxquels se rajoutent quatre acides aminés en C-terminal. Dans la structure de MAp19, affinée à 2.5 Å de résolution, nous observons globalement la même formation de dimère « tête-bêche » que dans C1s (figures D.1.1 et C.3.1), les ions calcium jouant le même rôle de stabilisation au centre de l'interface dimérique étendue, similaire à celle de C1s. Cependant, on peut noter une légère variation d'orientation au niveau de la charnière CUB-EGF entre les différentes protéines. Nous avons retrouvé dans la structure de MAp19 le deuxième site de calcium identifié précédemment dans C1s (site II), qui stabilise la partie distale du module CUB1.

D'un point de vue fonctionnel, nous avons pu identifier plusieurs acides aminés du site de fixation de la protéine MAp19 aux tiges de type collagène de la ficoline et de la MBL : les mutations des acides aminés présentés en rouge sur la figure D.1.2 ont conduit soit à l'absence totale de fixation de la protéine MAp19, soit à une diminution très importante des constantes d'affinité. Nous avons identifié un site commun de fixation à la MBL et aux ficolines L et H puisque les mutations affectent les liaisons à toutes ces protéines. Il se trouve que ce site de fixation se situe justement à proximité du site II de liaison du calcium dans le module CUB1 (publi n° 13).

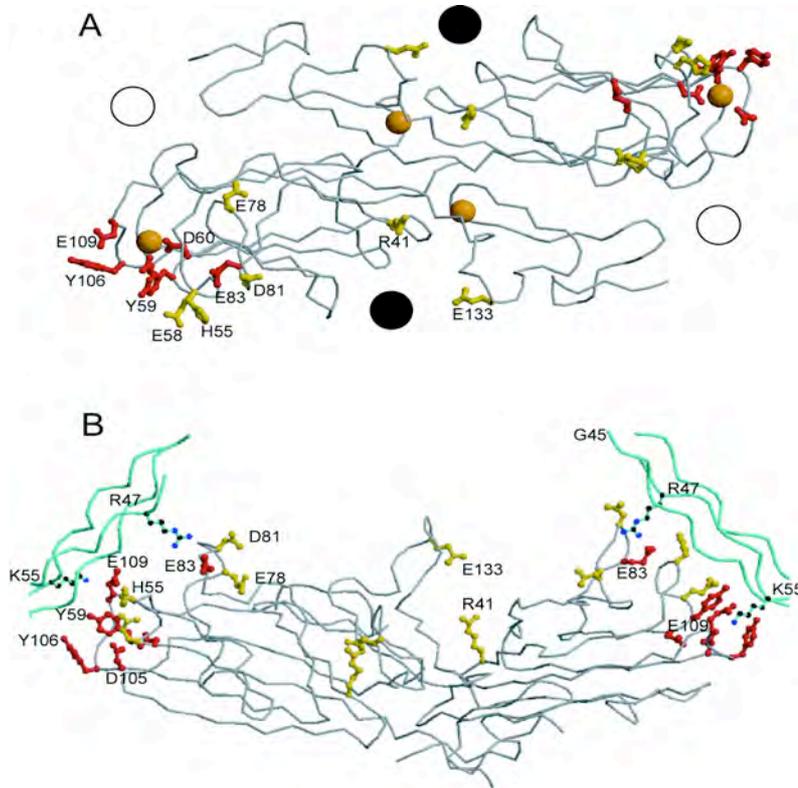
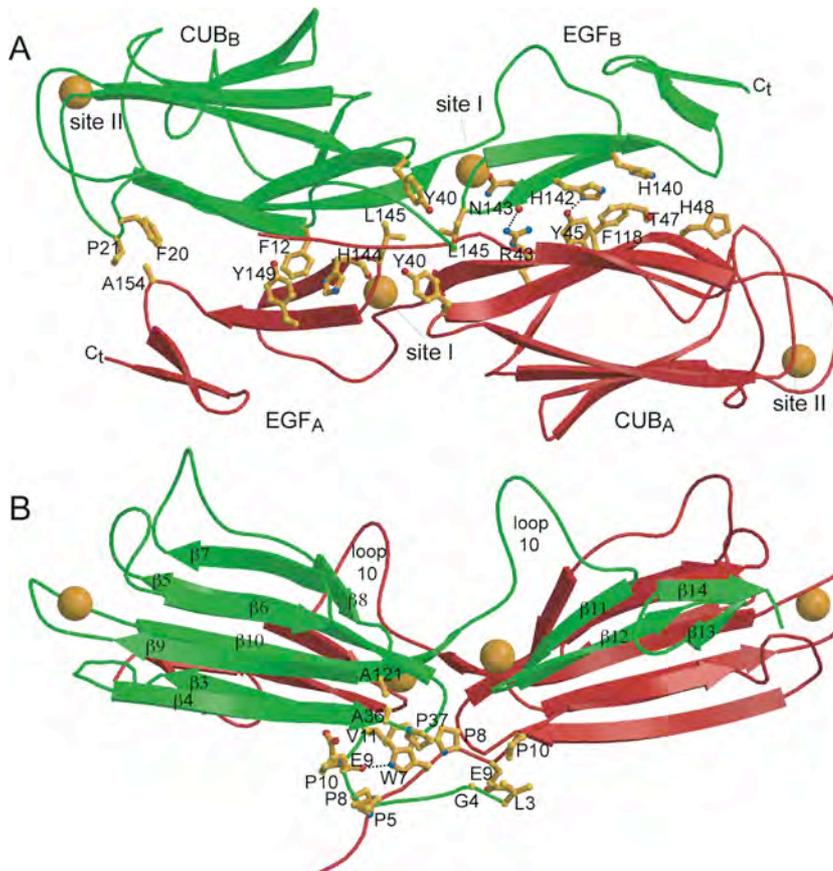


Figure D.1.2. Site commun de fixation de la MAP19 à la ficoline et à la MBL, défini par mutagenèse dirigée. Les acides aminés mutés sans incidence sur la fixation sont notés en jaune, alors qu'en rouge sont représentés les résidus dont la mutation entraîne soit la perte totale (D105G, Y59A, Y106A), soit une diminution très sensible de la fixation (D60A, E109A, E83A). (A) vue de dessus. Les cercles indiquent les sites de liaison proposés par Feinberg et al. (2003) (B) vue de côté, incluant un modèle d'interaction avec deux triples hélices de la MBL (cyan). (*publi n° 13*)

Une mutation naturelle a été identifiée dans le gène de la MASP2 chez un patient souffrant d'infections pulmonaires chroniques (Stengaard-Pedersen et al., 2003). Nos résultats permettent d'expliquer directement les conséquences de cette mutation (D105G), puisqu'elle affecte directement un des ligands du calcium dans le module CUB1 (publi n° 13), ce qui empêche l'interaction normale avec la MBL ou les ficolines. On peut noter à cette occasion que l'étude structurale des domaines CUB-EGF-CUB de la protéine MASP2 de rat, publiée précédemment par nos concurrents (Feinberg et al., 2003), ne permettait pas d'expliquer l'effet de cette mutation, puisqu'elle correspondait à une partie désordonnée de leur modèle.

D.1.2. Structure du fragment CUB-EGF-CUB de MASP1/3.

La structure du domaine CUB-EGF-CUB de MASP3 a aussi été résolue. Nous avons confirmé ainsi la présence de sites de fixation du calcium à l'extrémité du module CUB1, mais aussi dans le module CUB2. La caractérisation de mutants ponctuels a permis de mettre en évidence des résidus impliqués dans l'interaction MBL/MASP1, à proximité de chacun des sites de fixation du calcium dans les modules CUB1 et CUB2. Ceci nous permet de mieux comprendre l'influence du module CUB2 dans l'augmentation de l'affinité pour la MBL. Ces informations vont nous aider également à proposer un modèle d'assemblage entre la MBL et le domaine d'interaction des MASPs.

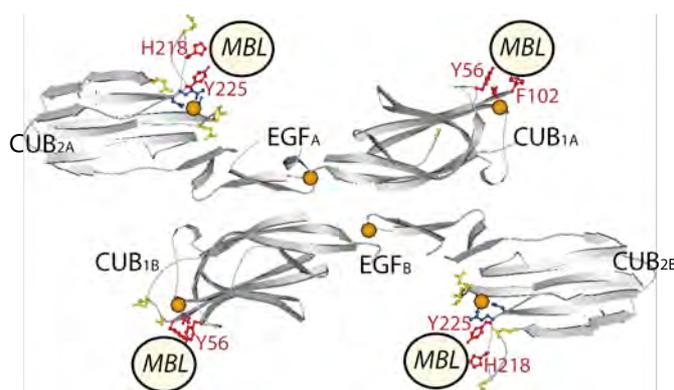


Figure D.1.3. MASP3, structure du fragment CUB-EGF-CUB et effets de mutations ponctuelles sur l'assemblage avec la MBL. Des mutations qui inhibent fortement la fixation de la MBL ou des ficolines sont marquées en rouge. Elles permettent de localiser 4 sites d'interaction.

D.1.3. Vers un modèle d'interaction MBL/MASPs.

Nous pouvons proposer un modèle d'association entre la MASP19 et les tiges de type collagène de la MBL (figure D.1.3). En plus des informations obtenues sur le site de fixation dans la protéine MASP19, la mutagenèse dirigée a permis de définir clairement un acide aminé du site de fixation de la MBL pour les MASPs : la Lys 55 (Teillet *et al.*, 2007).

Le modèle obtenu est suggéré sur la figure D.1.3, et présenté plus concrètement sur la figure D.1.4. Ce modèle diffère très fortement de celui proposé précédemment pour l'association du dimère C1r-C1s aux tiges de type collagène de C1q : ici le dimère se trouve à l'intérieur du cône défini par les tiges de type collagène (figure D.1.4), alors que, dans le cas du tétramère C1s-C1r-C1r-C1s, la fixation se produit sur la face externe (par rapport au cône) des tiges de type collagène de C1q (Figures C.1.2 et C.3.2). Il est ainsi paradoxal que les assemblages moléculaires observés entre la protéine de reconnaissance et les protéases associées apparaissent si différents dans les deux cas, alors que les deux types de complexe possèdent une architecture modulaire similaire à la fois du côté des protéases à sérine (modules CUB-EGF) et du côté des protéines de

reconnaissance (tiges de type collagène). Cela implique donc des différences subtiles dans les mécanismes moléculaires mis en jeu pour activer la voie classique et la voie lectine du complément. Néanmoins, les positions éloignées des domaines de reconnaissance, des domaines d'interaction et des domaines catalytiques implique que le signal d'activation sera aussi vraisemblablement transmis par un effet mécanique de déformation au contact de la surface cible. La synergie entre les 4 sites de fixation disponibles doit permettre de maintenir les protéases sur les tiges dans un contexte de stress mécanique.

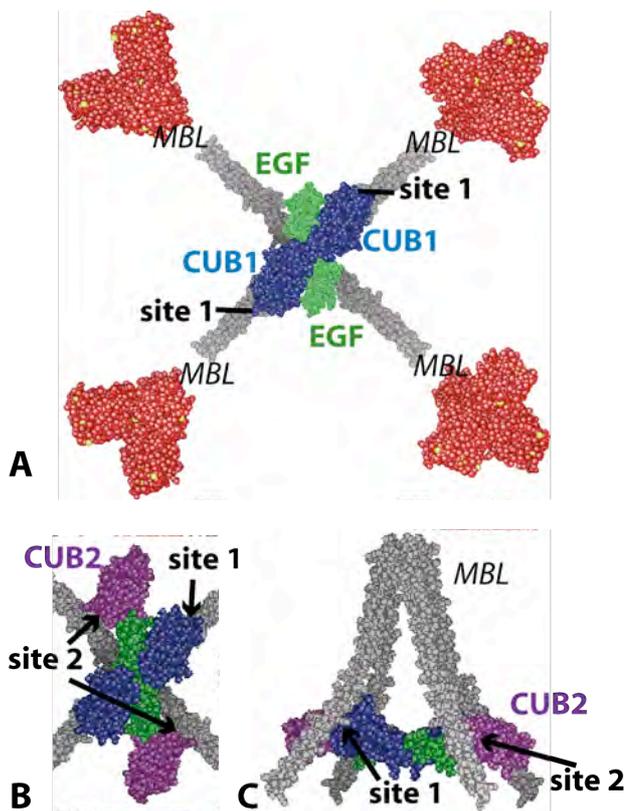


Figure D.1.4 .
Modèles d'association MBL/MASPs, (A) Modèle d'interaction MBL/MAp19 (vue de dessous). Les modules CUB1 et EGF de MAp19 sont notés, ainsi que la position du site 1 (d'après la publi n°12). (B,C) Vues de dessous et latérale du cône des tiges de MBL (gris) avec le fragment CUB-EGF-CUB de MASP1/3. Les 2 types de sites sont montrés. Le site 2 du module CUB2 renforce clairement l'interaction avec le domaine N-terminal des protéases (bleu-vert-magenta), au centre du complexe. Les domaines globulaires de la MBL sont en bas sur la figure C.

D.2. Domaine de reconnaissance des ficolines humaines.

L'objectif essentiel était de découvrir précisément par quelles interactions moléculaires les domaines de type fibrinogène des ficolines parviennent à reconnaître leur(s) cible(s). Pour cela, nous avons effectué une étude structurale des domaines de reconnaissance des ficolines en présence de ligands ou de molécules représentant une version simplifiée d'un ligand. Ces analyses cristallographiques ont nécessité de réaliser de très nombreux enregistrements de données de diffraction et de nombreux affinements de structure : 15 structures affinées dans la publi n° 10 sur les ficolines L et H, 5 structures affinées dans la publi sur la ficoline M en cours de soumission. Ces affinements ont été réalisés essentiellement par Virginie Garlatti dans le cadre de son travail de thèse, ainsi que par Nicolas Belloy lors de son stage M2.

D.2.1. Structure globale des domaines globulaires des ficolines.

Une structure trimérique globalement similaire à celle de la ficoline L (Figure D.2.1) a été obtenue pour toutes les ficolines: les fragments recombinants correspondant au domaine de type fibrinogène des ficolines s'associent spontanément en trimères.

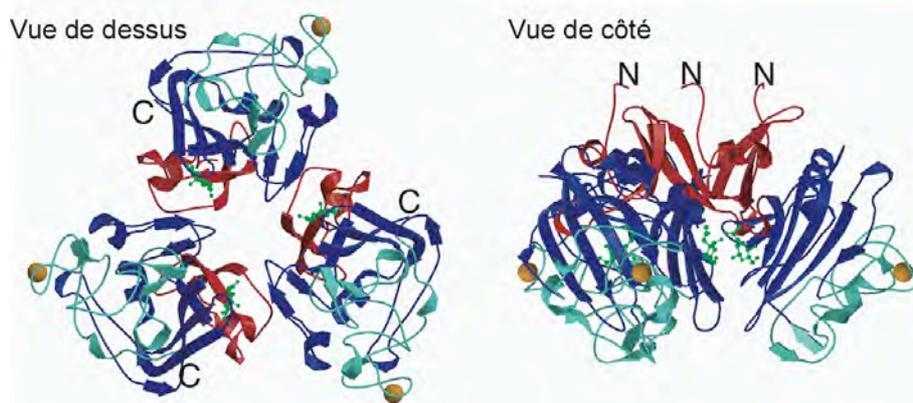


Figure D.2.1. Tête globulaire homotrimérique de ficoline L.

La structure du trimère est présentée sur la figure D.2.1. On peut remarquer que les trois extrémités N-terminales émergent du même côté du trimère, que l'on appellera la base du trimère. Leur conformation varie d'un monomère à l'autre, ce qui traduit sans doute le fait qu'il manque un élément de stabilisation: dans la molécule entière, ces extrémités N-terminales émergent à partir d'une triple hélice de type collagène. Le code de couleur utilisé sur la figure montre les trois sous-domaines A (rouge), B (bleu foncé) et P (cyan) composant les modules structuraux de type fibrinogène. Le domaine P est en position la plus externe, formant le « sommet » de la tête globulaire. Il comprend le site de liaison du calcium (sphère jaune). C'est aussi dans ce domaine qu'est attendu le site de liaison du ligand, à la fois en raison de sa position externe et par homologie avec le fibrinogène (site de polymérisation, d'où le nom P) et la tachylectine 5A de *Tachypleus tridentatus* (site de liaison du GlcNAc ; Kairies et al., 2001). On peut observer, sur la vue de côté de la tête globulaire de la ficoline L, une échancrure centrale dans la partie supérieure du trimère.

Si les trois ficolines présentent des associations trimériques similaires, des différences apparaissent dans la définition de leurs spécificités fines de reconnaissance, qui proviennent de substitutions d'acides aminés localisés en surface.

D.2.2. Un seul site de liaison externe observé pour les ficolines H et M.

Les ligands identifiés pour les ficolines H et M sont encore relativement peu nombreux. En corollaire, dans les deux cas, l'analyse structurale n'a révélé qu'un seul site de fixation externe. Ce site S1 est situé à proximité du site de liaison du calcium, dans le domaine P externe. Il correspond au site 'attendu', globalement similaire au site de fixation identifié précédemment dans la structure de la tachylectine TL5A d'un invertébré, la limule (Figure D.2.2). Cette similarité est plus marquée dans le cas de la ficoline M. La figure D.2.2 montre les principaux points communs, mais aussi les modifications subtiles d'acides aminés entre ces différents sites, qui modulent leur spécificités.

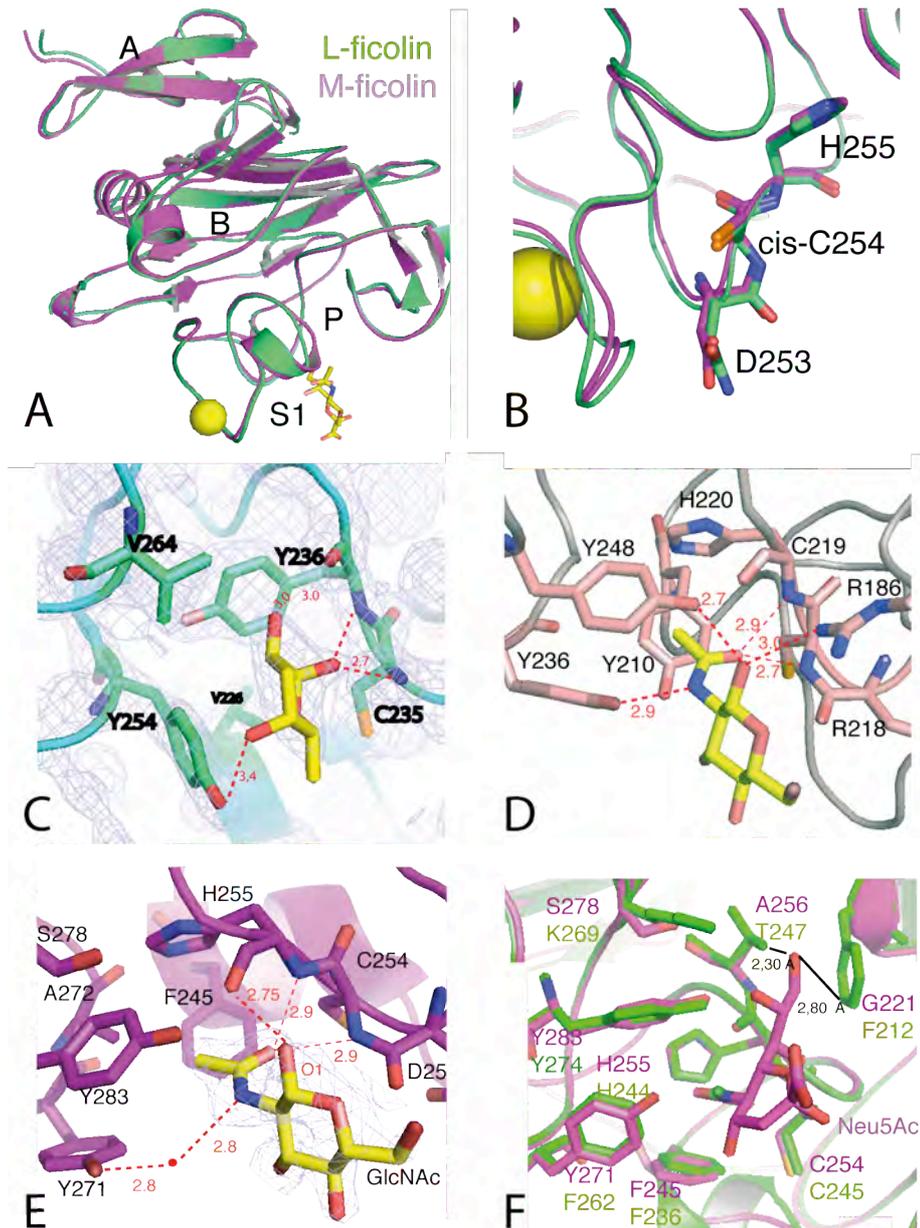


Figure D.2.2. Le site S1 de fixation du ligand dans les ficolines. (A) Superposition des ficolines L et M, montrant leur similarité et les positions du site S1 (le ligand de la ficoline M est montré en jaune) et du site de fixation du calcium (sphère jaune). Les domaines A, B et P sont aussi indiqués. (B) Zoom sur la cystéine liée par une liaison peptidique dans la conformation minoritaire *cis*. Cette conformation particulière est retrouvée dans toutes les ficolines, dans la tachylectine et dans le fibrinogène. (C) Site S1 dans la ficoline H, montrant les principaux éléments communs : la poche hydrophobe, au centre et à gauche, la cystéine (235) en position particulière, à droite, ce qui place les deux azotes consécutifs en position optimale pour stabiliser le ligand par des liaisons H, complémentaires de celle fournie par la tyrosine (254) sur le bord opposé. (D, E) GlcNAc dans les sites similaires de la tachylectine 5A et de la ficoline M. (F) Superposition montrant les variations du site S1 entre les ficolines L (vert) et M (magenta), avec son ligand Neu5Ac (magenta). La ficoline L ne peut accommoder ce ligand, faute de place. La phenylalanine 262, qui remplace la tyrosine 271 sur le bord gauche, ne peut plus stabiliser le ligand par liaison H.

D.2.3. Découverte d'un site de reconnaissance étendu pour la ficoline L.

L'étude structurale la plus fournie a été réalisée sur la ficoline L, plus étudiée, pour laquelle des spécificités de reconnaissance assez variées, voire contradictoires, ont été mises en évidence. En effet, une fixation de dérivés acétylés très variés avait laissé suggérer que cette molécule n'était pas vraiment une lectine (Krarup et al., 2004). Cependant, cette activité de type lectine était corroborée par la fixation spécifique de la ficoline L sur un sucre neutre étendu, le β -D-1-3-glucan, marqueur de la surface des levures (Ma et al., 2004). De façon inattendue, notre étude a permis de mettre en évidence dans la structure de la ficoline L des sites de fixation différents du site S1. L'ensemble des nouveaux sites observés (S2, S3, S4) définit une surface de reconnaissance étendue (Figure D.2.3).

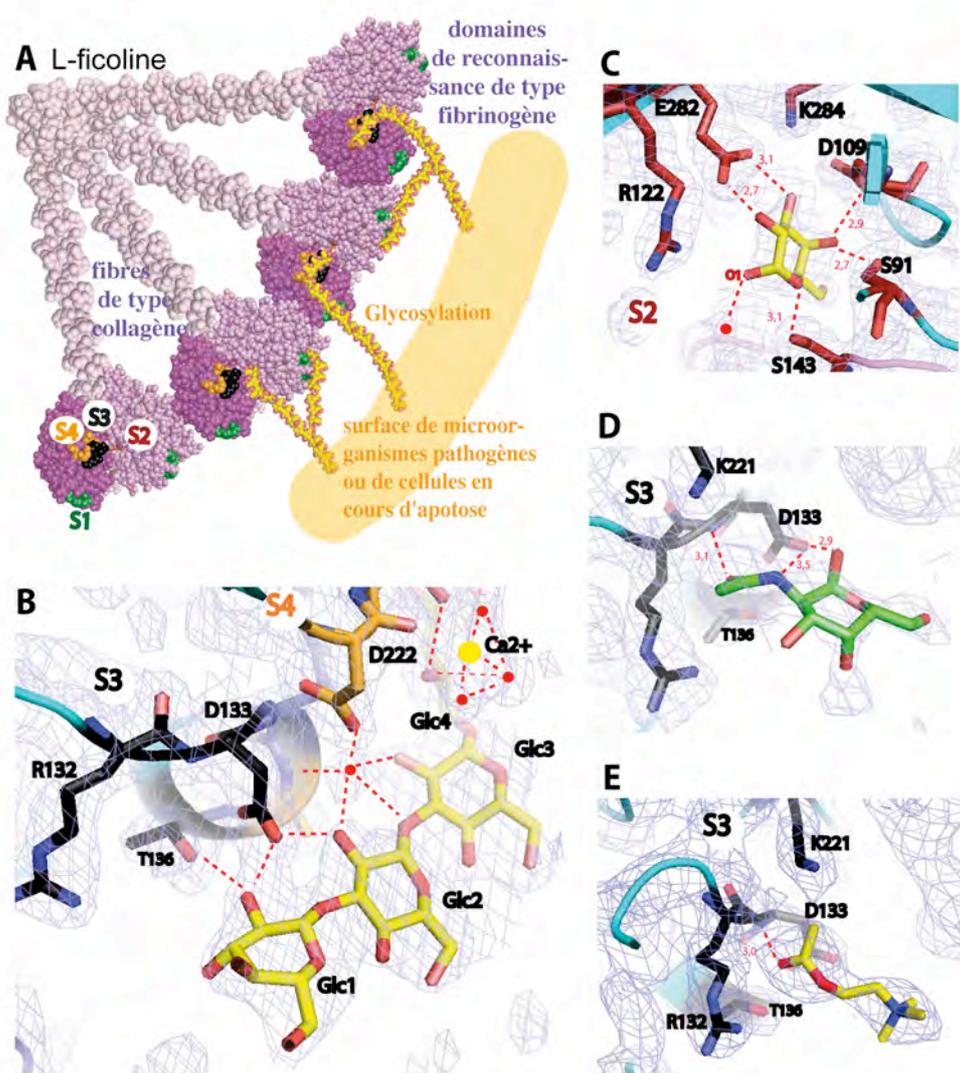


Figure D.2.3. Nouveaux sites identifiés dans la ficoline L (A) Représentation schématique de la ficoline L en interaction avec une surface. La position des sites S1 à S4 est montrée sur un des domaines de reconnaissance. (B) Un fragment de β -glucan (jaune) interagit avec les sites S3 (noir) et S4 (orange). (C) Une molécule de galactose (jaune) est stabilisée par de nombreuses liaisons-H dans le site interne S2 (rouge). (D, E) Divers ligands acétylés peuvent se lier au site S3 (noir) dans des orientations variées, comme ici le N-acétylgalactosamine (vert) et l'acétylcholine (jaune). La carte de densité électronique est beaucoup plus diffuse autour du sucre acétylé (D), qui présente des éléments de flexibilité.

Dans notre étude, deux ligands glucidiques neutres sont stabilisés par des réseaux de liaisons-H, ce qui confirme une activité de type lectine pour cette ficoline. En effet, tous les hydroxyles du galactose sont stabilisés dans le site interne S2 (Figure D.2.3.C), sauf en O1, qui est impliqué dans la liaison au sucre précédent dans des molécules plus longues. La liaison du β -glucan sur la partie externe des sites S3 et S4 est également remarquable (Figure D.2.3.B). Dans la petite poche du site S3, différents dérivés acétylés simples peuvent aussi être reconnus par leur seul groupement acétyle, dans des orientations variées (Figure D.2.3.D et D.2.3.E). Cette observation permet finalement de réconcilier les résultats biochimiques contradictoires évoqués précédemment, la ficoline L semblant avoir une spécificité de reconnaissance pour des molécules étendues, dont les éléments de reconnaissance peuvent inclure à la fois des motifs acétylés et des motifs de sucre neutre.

Cette propriété de reconnaissance de fragments moléculaires étendus, et non simplement d'un sucre terminal, est une perception nouvelle apportée par cette étude structurale. Elle aide à mieux comprendre la fixation spécifique d'autres ligands étendus par la ficoline L, comme les acides lipothéichoïques (Lynch et al., 2004) et les capsules des streptocoques de sérotypes III, groupe B (Aoyagi et al., 2005). Les acides lipothéichoïques (LTA : lipothéichoic acids) sont des glycolipides complexes, présents spécifiquement à la surface des bactéries gram positives (en particulier les Streptocoques et Staphylocoques). Dans ces deux types de ligands complexes, on peut noter la présence de GlcNAc, mais pas en position terminale, ce qui doit contribuer à leur reconnaissance par la ficoline L. Enfin nous avons perçu une relative plasticité du trimère de la ficoline L au cours de notre analyse structurale.

D.2.4. Particularités structurales du site S1 de la ficoline M.

Les interactions observées dans le site S1 de la ficoline M expliquent sa spécificité de reconnaissance pour les composés acétylés comme le GlcNAc (Figure D.2.2.E), le GalNAc. Une particularité de la ficoline M est de pouvoir interagir aussi avec le Neu5Ac (Figure D.2.2.D), et les acides sialiques de façon plus générale. Cette reconnaissance additionnelle des acides sialiques est limitée à cette ficoline M, et à son homologue chez la souris, ces deux ficolines étant sécrétées par des cellules de la lignée monocyttaire. Elle n'existe pas pour les autres ficolines plasmatiques, produites par le foie. Cette spécificité peut expliquer la détection initiale de cette protéine à la surface des monocytes, car les cellules immunes contiennent des fortes concentrations d'acides sialiques sur leur surface (Crocker et al, 2007). En effet, les acides sialiques sont connus essentiellement pour être des marqueurs du « soi », car ils sont présents à la surface de nombreuses molécules glycosylées. Néanmoins, un certain nombre de pathogènes présentent des acides sialiques à leur surface, ceci sans doute dans une stratégie de mimétisme de la surface des cellules de l'hôte. On peut citer par exemple *Neisseria meningitidis*, *Campylobacter jejuni*, *Trypanosoma cruzi* et les capsules du groupe B de Streptocoques (Crocker et al, 2007).

Une autre particularité observée lors de l'étude structurale de la ficoline M est le changement de conformation des boucles autour du site S1 à pH acide (pH 5.6). Le site S1 est complètement détruit, les éléments formant le site subissant des déplacements importants dans des directions différentes. Ce changement de conformation est très probablement induit par le changement de charge de plusieurs résidus histidine dans cette partie de la molécule. En effet, la comparaison des structures permet de repérer trois histidines qui sont très bien stabilisées à pH neutre par des liaisons H, cette stabilisation étant perdue dans la conformation observée à pH acide.

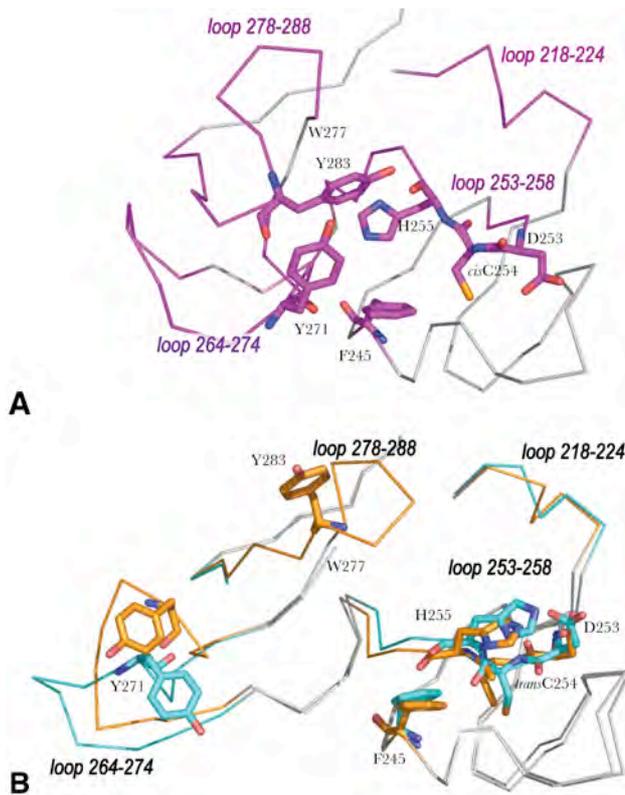


Figure D.2.4. Changement de conformation dépendant du pH dans la ficoline M.

A) Regroupement des acides aminés du site S1 à pH neutre. B) A pH acide, on note un changement de conformation important des boucles 264-274, 278-288 et 253-258. La plupart des résidus formant le site S1 à pH neutre sont portés par ces boucles, et se retrouvent à pH acide en disposition ‘éclatée’ non fonctionnelle. (publication en cours de soumission)

Les effets de ce changement de conformation étant remarquablement ciblés sur le site S1, il est tentant de supposer qu’il peut jouer un rôle fonctionnel, induisant par exemple la dissociation du complexe ficoline-ligand à pH acide. Parmi les compartiments physiologiques présentant un pH acide, on peut citer les phago-lysosomes. Il se trouve que la ficoline B de souris a été récemment détectée au niveau des phagolysosomes de souris (Runza et al., 2006). Or la ficoline B est l’homologue de la ficoline M chez la souris, présentant également les particularités de ne pas être détectée dans le plasma, d’être sécrétée par les cellules de la lignée monocyttaire et de se fixer aux composés sialylés (Endo et al., 2005). Même si le rôle fonctionnel de ce ‘switch’ conformationnel apparaît très probable, cette hypothèse devra être confirmée expérimentalement.

D3. Petite conclusion sur l’étude de la voie lectine.

Nous avons réussi à nous introduire sur le terrain concurrentiel de l’étude structurale des complexes initiant la ‘voie lectine’ du complément. En particulier, nous avons fourni à la communauté scientifique des résultats de qualité qui permettent de mieux comprendre les effets de mutations naturelles de la protéine MASP2 rencontrées chez certains patients (publi n°13) ou de commencer à déchiffrer différents modes de reconnaissance spécifique des ficolines. Par rapport aux études structurales effectuées précédemment sur la voie classique, le schéma fonctionnel final des complexes

MBL/MASPs ou ficolines /MASPs est plus flou, mais clairement distinct de celui identifié pour C1. Les variations de composition des différents complexes compliquent un peu cette analyse, et laissent au moins suggérer un rôle de régulation aux protéases MASP1 et MASP3, dont la fonction reste à éclaircir.

Par rapport aux études précédentes sur C1, les études structurales se sont enrichies ici d'expériences de mutagenèse dirigée pour identifier les sites d'interaction MBL/MASPs (publi n° 13 et publication en cours de préparation), ou d'études cristallographiques de nombreux complexes avec des ligands variés pour l'analyse des spécificités des ficolines (publi n°10). L'identification d'un nouveau site étendu de fixation des ligands dans la ficoline L a permis de mieux comprendre les propriétés de reconnaissance controversées de cette molécule, et ouvre la voie vers de nouvelles études structurales en complexe avec des ligands physiologiques plus complexes et étendus comme les acides lipotéichoïques ou les capsules polysaccharidiques de Streptocoques de type B. Parmi les nouveaux ligands identifiés, la fixation de galactose, dans une orientation compatible avec la liaison d'un galactose en position terminale, permet de formuler des hypothèses nouvelles de motifs moléculaires possibles pour la reconnaissance des cellules apoptotiques. De même, la reconnaissance spécifique de D-fucose par la ficoline H est compatible avec son spectre de reconnaissance étroit et permet de suggérer quelques micro-organismes potentiellement reconnaissables. Enfin, nous avons pu analyser également, dans le cas de la ficoline M, un phénomène de changement de conformation dépendant du pH (publication en cours de soumission), suggérant un nouveau mécanisme fonctionnel.

E. Aperçu technique expérimental des projets.

Pour compléter la présentation de la problématique scientifique, j'ai regroupé ici quelques informations sur la réalisation pratique et technique des projets. Le cheminement expérimental de l'analyse cristallographique d'une protéine ou d'un fragment protéique est représenté schématiquement sur la figure E.1. Même si ce processus paraît linéaire, des difficultés peuvent apparaître à différentes étapes, nécessitant éventuellement un retour aux étapes précédentes, voire un abandon du projet. Chacune des études cristallographiques réalisées (14 au total) a présenté des difficultés qu'il a fallu surmonter. La nature de la difficulté et l'étape concernée a été variable d'un projet à l'autre. C'est cet aspect technique de l'approche expérimentale cristallographique que je vais très rapidement résumer dans ce chapitre.

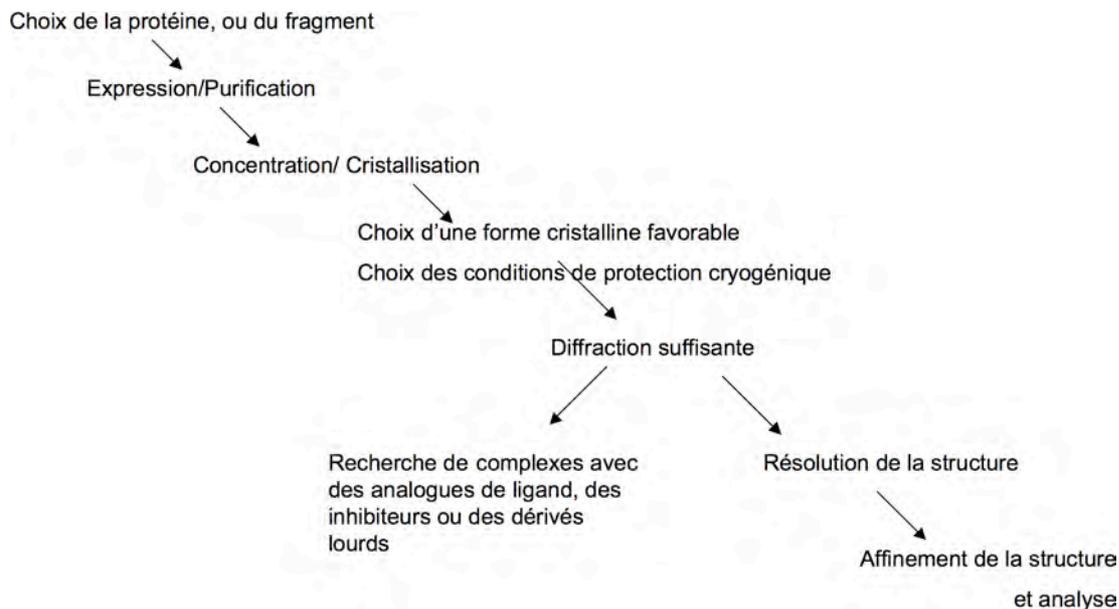


Figure E.1. Schéma global des principales étapes d'une étude cristallographique

Les étapes initiales de production, purification, concentration et cristallisation, telles que nous les avons utilisées, dépendent essentiellement des propriétés biochimiques uniques à chaque protéine.

E.1. Production-Purification

Les premières étapes, jusqu'à la purification, ont été réalisées par mes collaborateurs, dans le groupe de Gérard Arlaud. Quelques fragments recombinants n'ont pas pu être produits en quantité suffisante pour les études cristallographiques, comme le fragment N-terminal de C1r. Le système d'expression qui a été privilégié pour ces études est le système baculovirus et cellules d'insectes, dans des conditions où les protéines sont exprimées sous forme soluble dans le surnageant de culture. Ce système présente l'avantage de produire des protéines correctement repliées, avec synthèse correcte des ponts disulfures, et avec un certain nombre de modifications post-traductionnelles relativement proches de celles effectuées chez les eucaryotes, comme par exemple les glycosylations. Ce système ne convient pas pour la production de fibres de type collagène pour lesquelles les modifications post-traductionnelles sont nombreuses et atypiques. Il ne convient pas non plus pour l'expression des hétérotrimères des têtes globulaires de C1q. Enfin, ce système n'est pas particulièrement adapté à la production de protéines sélénées.

E.2. Concentration :

La première étape où nous pouvons être confronté à des difficultés techniques est la concentration. En effet, la cristallisation des protéines demande non seulement un échantillon de protéine très pur, mais aussi très concentré. Même si de nombreux projets d'études cristallographiques utilisent des concentrations en protéine supérieures à la dizaine de mg/ml, nous avons utilisé dans nos études des concentrations plus faibles (Tableau E.1), parfois difficiles à atteindre. L'utilisation d'un agent solubilisant non détergent (Vuillard et al., 1994; Vuillard et al., 1995), la sulfobétaïne 195 (NDSB195), nous a aidé sur plusieurs projets à dépasser le seuil du mg/ml (dont 2 projets avec succès, voir tableau E.1).

Tableau E.1 Exemples de concentrations et conditions de cristallisation utilisées dans différents projets.

Protéine/fragment	Concentration utilisée Conditions de cristallisation
Trypsine humaine	41 mg/ml 14-16% PEG 6K, dioxane 2-3%
C1s CCP2-SP	6 mg/ml PEG4K 32%, pH 5.7 (MES), 4°C PEG4K34%, pH7.4 (Hepes), 100mM sulfate d'ammonium, 4°C
C1r CCP1 CCP2-SP	2-6 mg/ml sulfate d'ammonium 1.5 M, pH 8.5 (TAPS)
C1r CCP2-SP mutant S637A	8 mg/ml sulfate d'ammonium 1.5 M; pH 8.4 (TAPS)
C1r CCP2-SP activé	6 mg/ml PEG6K 12-20%, pH 7.5 (Hepes), addit. 20% glycérol ou 5% PEG400
gC1q	3-5 mg/ml 0.1M NDSB195 2% glycerol PEG4K 37-41%, pH 7.0 7.2 7.4 (Tris-HCl) additifs 50mM CaCl ₂ , 10mM β-mercaptoethanol
C1s CUB-EGF	6-7.5 mg/ml PEG4K 16%, 0.2M MgCl ₂ , addit. pH 7.5 (Hepes), 8.7 % glycérol Germe cristallins: PEG400 30%, CaCl ₂ 0.2M, pH 7.5 (Hepes)
MAp19	2.5 mg/ml 3% PEG8K, pH 8.5 (Tris-HCl)
MASP3 CUB-EGF-CUB	3.6 mg/ml PEG8K 18-20% Hepes pH 7.0 glycerol 3%
Ficoline L	5.7 mg/ml PEG 8K, pH 7.0 (Hepes), 0.2 M acétate de calcium. Ensemencement. Na succinate 0.9-0.95 M, pH 6.5 or pH 7.0
Ficoline H	3 mg/ml 0.1 M NDSB195 PEG 6K 21-26%, pH 8-8.5 (Tris-HCl), 20mM CaCl ₂
Ficoline M	6 mg/ml PEG 4K 11%, addit. isopropanol 5%, pH 7.0 (Hepes) PEG 4K 23%, addit. LiSO ₄ 0.3 M, pH 5.6 (MES)
MASP3 CUB2-CCP1	2mg/ml NDSB195 0.1 M sulfate d'ammonium

E.3. Cristallisation

La cristallisation d'une protéine, c'est-à-dire son empilement ordonné selon les trois directions de l'espace en respectant la symétrie d'un cristal, comporte une composante

aléatoire, pour ne pas dire un aspect ‘miraculeux’. Pour atteindre cet état solide ordonné très particulier, il faut diminuer la solubilité de l'échantillon jusqu'à se trouver dans des conditions thermodynamiques (et cinétiques) favorables à la nucléation puis à la croissance d'un cristal. De très nombreux paramètres permettent de jouer sur la solubilité d'une protéine, et l'ensemble des combinaisons possibles ne peut pas être exploré de façon systématique. Cette étape peut être limitante pour certains projets.

La tendance actuelle est, dans un premier temps, de cribler avec des robots de cristallisation de nombreuses conditions prédéfinies, disponibles dans des kits commerciaux. Le robot disponible sur la plateforme de biologie structurale voisine (PSB) permet de faire des gouttes de très faible volume, de l'ordre de 100 nl. Si on a la chance de trouver à cette étape une ou plusieurs conditions qui donnent des résultats positifs (oursins, microcristaux, agrégats de cristaux, aiguilles, ...), il faudra alors chercher ensuite à les améliorer. Avec vraiment beaucoup de chance, on peut même obtenir directement de jolis petits cristaux, que l'on pourra reproduire et légèrement améliorer.

Les échecs rencontrés à cette étape peuvent se classer en trois catégories principales:

a) pas de cristaux de taille suffisante obtenus de façon reproductible.

Ceci nous est arrivé pour plusieurs fragments de MASP-3, à l'exception du fragment N-terminal CUB-EGF-CUB. Cette protéine ayant de nombreux sites de glycosylation, des fragments pour lesquels certains sites ont été mutés ont été construits.

b) des cristaux, mais pas le contenu recherché.

Dans ce cas, on arrive bien à obtenir des cristaux, mais il ne contient pas ce qui est étudié. Cela peut être juste du sel, ou juste la protéine d'intérêt sans son partenaire. Ceci nous est arrivé pour de nombreux essais de co-cristallisation : par exemple les têtes globulaires de C1q avec des ligands ou avec le récepteur gC1qR, ou pour les ficolines avec des ligands. Pour les ligands de petite taille, il est possible d'utiliser la technique alternative de trempage des cristaux de protéine dans une solution contenant le ligand (voir E.5).

c) des cristaux du fragment étudié aux qualités de diffraction limitantes:

Le dernier critère à respecter est évidemment la qualité de diffraction des cristaux, sans laquelle il n'est pas possible d'obtenir les informations expérimentales nécessaires. Les progrès techniques réalisés au niveau des lignes de lumière du synchrotron voisin (ESRF) ont beaucoup amélioré le succès à cette étape, mais il reste des projets pour lesquels cela reste un facteur limitant. Lorsque de nombreuses conditions de cristallisation sont obtenues, il est indispensable de les sélectionner ensuite suivant le critère de la diffraction aux rayons X. Dans le cas de la **ficoline M**, par exemple, certaines conditions donnaient des cristaux qui diffractaient très bien (mieux que 2 Å, 1.5 Å) et d'autres de façon presque insuffisante (3 Å au mieux).

Pour l'étude cristallographique de la **ficoline L**, nous avons obtenu également plusieurs formes cristallines, dont certaines étaient obtenues dans les mêmes conditions de cristallisation. Les plus beaux cristaux étaient souvent caractérisés par une maille qui n'était pas très favorable à l'étude cristallographique. En effet, la maille obtenue était très grande pour une molécule de cette taille (forme D, groupe d'espace P2₁2₁2, a=109 Å, b=130 Å, c=270 Å), car elle contenait 4 trimères par unité asymétrique, ce qui veut dire 12 sous-unités à prendre en compte lors de l'étape d'affinement. De plus, la qualité de diffraction de ces cristaux était assez limitée (2.6 -2.8 Å). Cette forme cristalline nous a permis de confirmer de façon qualitative la position de certains ligands dans cet empilement cristallin, mais nous avons utilisé d'autres formes cristallines pour effectuer ce projet plus efficacement.

Pour l'étude cristallographique du **fragment CCP1-CCP2-SP de C1r** sous forme activée (figure C.4.1), une seule forme cristalline a été obtenue, qui contenait bien le fragment étudié. Les cristaux étaient magnifiques, de grande taille, en forme de diamants. Cependant ils contenaient un fort pourcentage de solvant (>80%), ce qui limite souvent la qualité de diffraction des rayons X. De plus, ils étaient caractérisés par une grande maille ($a=b=101.8 \text{ \AA}$, $c=461.6 \text{ \AA}$, groupe d'espace $P6_1$), ce qui limitait la distance cristal-détecteur utilisable pour garder les tâches de diffraction séparées. Un jeu de données de diffraction de qualité moyenne (4 Å résolution) a pu être obtenu sur la ligne BM30 à l'ESRF. Celui-ci était de qualité insuffisante pour affiner un modèle, mais nous a permis de confirmer la forme globale de l'enveloppe du dimère (publi n° 17).

Pour le fragment CUB2-CCP1 de MASP3, la diffraction des tout petits cristaux obtenus de façon reproductible a été totalement insuffisante.

E.4. La résolution de la structure

Les informations de diffraction du cristal ne permettent pas de visualiser directement l'image de la molécule, ni sa carte de densité électronique. En effet, il manque une information cruciale, dite 'information de la phase', pour calculer la carte de densité électronique. Il est donc nécessaire de résoudre le problème de la phase pour déterminer et visualiser la structure, étape appelée de façon plus concise 'résolution de la structure'. La méthode qui a été le plus utilisée sur les projets présentés ici est appelée le 'remplacement moléculaire'. Elle consiste à utiliser un ou plusieurs modèles de fragment homologue et à en trouver l'orientation dans la maille de façon à obtenir une première estimation calculée de la phase. On procède ensuite de manière itérative pour améliorer le modèle, et donc les phases.

E.4.1. Structure du domaine catalytique de C1s :

Cette structure a été résolue par remplacement moléculaire, en utilisant les modèles de protéases à sérine de structure connue. Un premier jeu de données de diffraction a été enregistré sur la ligne BM30, à l'ESRF, à une résolution de 2.2 Å. Avec ce premier jeu de données de diffraction, il était impossible de reconstruire correctement la totalité du modèle. En effet, même si la solution de remplacement moléculaire était correcte, le nombre de données expérimentales de réflexions était insuffisant pour surmonter le biais introduit par le modèle de départ, en particulier pour toute la région de la protéase définissant le site actif et le site de reconnaissance du substrat, ces zones variant d'une protéase à l'autre et de l'état proenzyme à l'état activé. Un jeu de données de diffraction de très bonne qualité, jusqu'à 1.7 Å de résolution, obtenu sur la ligne ID14-EH4 de l'ESRF sur un autre cristal, a permis de reconstruire complètement le modèle sans ambiguïté. La procédure de reconstruction automatique arp-warp, au stade de développement à cette époque, a permis de reconstruire l'essentiel (92%) du modèle final à partir d'un modèle initial très incomplet (50 %). Cette procédure ne reconstruisait que la chaîne principale du modèle. La carte de densité électronique obtenue était cependant de très bonne qualité et l'identification des chaînes latérales s'est faite sans aucune ambiguïté. Nous avons alors constaté que le fragment catalytique CCP2-SP avait cristallisé sous la forme activée, même si le fragment recombinant avait été purifié sous la forme proenzyme et que les tests d'autoactivation à 37°C avaient été négatifs.

E.4.2. Structure du domaine globulaire de C1q :

La détermination expérimentale de la structure cristallographique des têtes globulaires de C1q a été un processus long (une dizaine d'années) et difficile. De nombreuses personnes ont contribué à l'avancée de ce projet, en particulier Claudine Darnault pour l'amélioration de la qualité des cristaux, Arnaud Gruez et David Pignol pour l'enregistrement de nombreux jeux

de données de diffraction au synchrotron et pour les premiers essais de résolution de la structure, Monique Lacroix pour la purification de nombreux lots de protéines. Je n'ai travaillé sur ce projet que dans les étapes finales de la résolution, de la reconstruction et de l'analyse de la structure, après avoir aidé à l'enregistrement de quelques jeux de données de type SAD ou MAD au synchrotron. J'ai réalisé l'étape cruciale de résolution de la structure par remplacement moléculaire et Jordi Juanhuix, en stage post-doctoral, m'a aidé pour construire la fin du modèle, affiner et analyser la structure finale.

La première difficulté à surmonter pour la résolution de la structure était d'identifier la nature des trois chaînes homologues des modules gC1qA, gC1qB et gC1qC qui composent le domaine globulaire. Pour cela, j'ai utilisé les informations obtenues sur un cristal dans lequel des atomes lourds avaient diffusé. La position de 2 atomes lourds de platine par trimère a été déterminée, le cristal contenant 2 trimères dans l'unité asymétrique (groupe d'espace P1). Sachant que les atomes de platine interagissent de façon préférentielle avec les groupements sulfurés, j'ai recherché la présence d'une éventuelle cystéine ou méthionine au voisinage des sites de platine identifiés. Un des sites était proche de la molécule, vers la position Gln142 du modèle initial (*). L'alignement de séquences (Figure E.2) montre un seul candidat possible: la méthionine de la chaîne B. L'autre site était plus distant, correspondant vraisemblablement à une région plus variable, et les éléments les plus proches étaient soit la position 137 ou 139 de la chaîne B (celle de Gln142), soit la molécule symétrique autour des positions 126 ou 128 (+). Là encore, une seule solution ressort clairement de l'alignement de séquences : la méthionine de la chaîne A.

Site proche (Gln142) :		Site lointain (vers 126-128)
	*	+ +
ACRP30	F Y N Q Q N H	T R V T V
C1q A	I T N Q E E P	N P P M
C1q B	I T N M N N N	
C1q C	L T N P Q G D	T H Q PPA

Figure E.2. Alignements de séquence utilisés pour identifier les chaînes de C1q.

Ce premier pari sur la nature des chaînes A, B et C a conduit à revoir complètement la numérotation des chaînes d'un trimère et a permis de connaître la séquence de chaque chaîne.

Le deuxième obstacle à surmonter était celui du biais du modèle, d'autant plus que le modèle de départ (adiponectine ACRP30) n'était pas de bonne qualité et possédait des boucles non définies. J'ai éliminé toutes les parties variables du module (variables dans les alignements de séquence), en ne gardant que le 'cœur' conservé de ces modules (fragments 116-118, 135-165, 176-179, 199-217, 234-244), pour lequel une corrélation locale de la densité électronique avec la densité calculée était supérieure à 0.65. Ce nettoyage des parties incorrectes du modèle a sensiblement amélioré les résultats de l'étape d'affinement en corps rigide de la position relative des modules (pseudo-symétrie d'ordre 3) à 3.5 Å, une étape qui ne converge pas facilement. Les calculs ont été réalisés en utilisant le programme d'affinement CNS, les valeurs des facteurs R de désaccord restant très élevés à cette étape (R 47.8, Rf 46.2 à 3.5 Å ; R 48.8, Rf 47.9 à 3 Å). Ensuite, il a fallu corriger la séquence des fragments présents dans ce modèle, et reprendre ensuite l'affinement en corps rigide, pour mieux définir les contraintes de symétrie non-cristallographique, afin de faciliter la convergence des affinements. En effet, la résolution du jeu de données du dérivé platine était limitée à 2.6 Å. Pour définir avec précision la matrice d'orientation définissant l'axe binaire non-cristallographique, j'ai utilisé le logiciel graphique O en superposant de façon réciproque les coeurs de trimère, non pas les sites de platine. Cela m'a permis d'améliorer encore légèrement ce modèle de départ (R 43.0, Rf 44.7 à 2.6 Å). Les essais rapides d'amélioration

de la carte de densité électronique par moyennation ou en utilisant le logiciel warp n'ont pas permis de progresser à cette étape.

Ce nouveau « modèle minimum » affiné a permis d'obtenir la bonne solution de remplacement moléculaire avec les données de diffraction enregistrées à plus haute résolution (1.85 Å) sur une forme cristalline différente (groupe d'espace C2), en gardant la bonne identification des chaînes A, B et C. En superposant différents types de cartes électroniques (3mFo-2DFc, warp), différents modèles entiers (adiponectine ACRP30 et domaine globulaire du collagène X) et partiels (fragments reconstruits automatiquement par arp-warp), il a été possible de reconstruire sans problème l'ensemble de la structure du domaine globulaire.

E.4.3. Structure du fragment CUB-EGF de C1s :

La détermination expérimentale de la structure du fragment CUB-EGF de C1s a été réalisée par Lynn Grégory au cours de son travail de thèse. La première étape a été de contrôler le processus de cristallisation pour obtenir de façon reproductible des cristaux uniques ayant un bon pouvoir de diffraction, en utilisant une technique assez originale d'ensemencement croisé. Comme dans le projet précédent, l'étape de résolution de la structure n'était pas très facile, le groupe d'espace étant triclinique (P1) et les différents jeux de données enregistrés ne respectant pas les contraintes d'isomorphisme. Plus de 27 composés d'atomes lourds ont été essayés, en testant différentes concentrations et temps de trempage.

Parmi ces essais, nous avons tenté de remplacer les ions calcium (attendus au niveau du module EGF) par des ions de la famille des lanthanides afin d'utiliser le fort signal anomal au seuil d'absorption LIII. Seul le chlorure de terbium a donné des résultats positifs (fixation de deux atomes), et différentes étapes d'optimisation des conditions de trempage ont été nécessaires (il fallait par exemple faire un pré-trempage pour diminuer la concentration initiale de magnésium dans les cristaux). Nous avons eu de la chance à cette étape, car le calcium n'a pas été substitué dans le site connu (module EGF) mais dans un deuxième site, inattendu mais plus accessible, dans chaque module CUB (site II, Figure C.3.1).

Deux jeux de données de diffraction ont été enregistrés à 3 Å de résolution (sur BM30 et BM14), au seuil d'absorption LIII du terbium, ce qui a permis d'obtenir une première estimation des phases par la méthode SAD (single anomalous dispersion). Les brins les plus visibles ont été tracés dans les premières cartes, après aplatissage du solvant. Le calcul d'une carte de densité électronique moyennée entre les deux monomères de l'unité asymétrique a permis de tracer les modules CUB, mais la qualité de la carte s'est alors dégradée au niveau des modules EGF. Un premier modèle a ainsi été progressivement obtenu. Même si ce modèle était encore très incomplet (60%), il a été affiné puis utilisé comme modèle de remplacement moléculaire avec le jeu de données à haute résolution (1.5 Å) pour finir de résoudre la structure. A cette étape, le programme arp-warp a été utilisé, ce qui a permis la construction automatique de 86% du modèle. La fin du modèle a été construite de nouveau manuellement.

E.4.4. Structure de la protéine MAp19

La structure de la protéine MAp19 a été résolue par remplacement moléculaire à partir de la structure du fragment CUB-EGF-CUB de la protéine MASP2 de rat (Feinberg et al., 2003). Même si l'on observe globalement la même formation de dimère « tête-bêche » que dans C1s, l'utilisation de la structure de C1s ne permettait pas de résoudre la structure par remplacement moléculaire. Si l'on compare précisément les dimères de C1s, MASP2 de rat et MAp19, on peut constater en effet des variations subtiles dans les positions relatives des modules CUB et EGF, à l'intérieur d'un même monomère (Figure E.4.1) ou à l'interface du dimère (publi n°13). Ceci montre qu'il y a une légère flexibilité au niveau de la charnière CUB-EGF, tout du moins lorsqu'on passe d'une protéine à l'autre. La technique de

remplacement moléculaire, par laquelle nous avons tenté de retrouver la position de paires de modules CUB/EGF, est très sensible à ces variations subtiles d'orientation. La recherche de la position des modules isolés est également rendue difficile par leur petite taille, le manque de structures secondaires, leur forme très allongée et leurs contacts étroits sur une large surface: on est bien loin du modèle sphérique pour lequel cette technique a été développée !



Figure. E.4.1. Superposition des modules CUB (à droite) et EGF (à gauche) d'un monomère de MAp19 (rouge) et de C1s (jaune). (vue stéréoscopique)

E.4.5. Brève conclusion sur la méthode du remplacement moléculaire.

J'ai accumulé une certaine expérience sur l'utilisation de cette méthode à travers les différentes études cristallographiques réalisées (14 au total). Parmi les motifs structuraux étudiés, les domaines de type protéases à sérine représentent vraiment un cas idéal, présentant une forme globulaire et contenant au moins une hélice. Au contraire, les petits modules de forme allongée, comportant très peu de structures secondaires (CCP, EGF) ou uniquement des brins (CUB, gC1q), ne sont pas vraiment adaptés, et nous n'avons jamais pu les utiliser seuls. Il est possible de les utiliser dans une version 'pré-assemblée' dans un certain nombre de cas favorables (CCP2-SP, trimère globulaire de gC1q) mais la méthode du remplacement moléculaire est très sensible aux décalages relatifs d'orientation (cf MAp19 et CUB-EGF de C1s). L'affinement de la position de ces modules par corps rigide a souvent aussi conduit à des erreurs, la plus fréquente étant une position décalée d'un brin sur un feuillet β (cas rencontré encore dernièrement lors de la résolution de la structure du fragment CUB-EGF-CUB de MASP3). Enfin, pour minimiser au maximum le biais introduit par le modèle de départ, nous avons utilisé de nombreuses fois la procédure arp-warp (Morris et al., 2003).

E.5. Diffusion de ligands dans des cristaux par trempage.

Il est parfois nécessaire de tester un bon nombre de conditions de trempage avant d'aboutir à la structure d'un complexe protéine-ligand. Cela est vrai en particulier dans le cas de sites de fixation d'affinité assez modérée, situation que nous rencontrons pour les protéines oligomériques que nous étudions, cette faible affinité étant compensée 'in vivo' par l'avidité de leur fixation multiple sur une surface. C'est la difficulté principale que nous avons eu à surmonter sur le projet d'étude cristallographique des ficolines. Les causes d'échec à cette étape peuvent être multiples :

- a) le cristal diffracte bien après trempage, mais il n'y a pas de densité électronique supplémentaire correspondant au ligand. -> il faut essayer d'augmenter les concentrations et/ou la durée du trempage
- b) Le cristal ne supporte pas bien le trempage (il 'craque') et/ou la diffraction n'est pas bonne -> il faut alors revoir le contenu de la solution de trempage, diminuer le temps de trempage, ou diminuer les concentrations.

Il n'y a pas de solution universelle. Nous avons utilisé de très fortes concentrations (500mM) de ligands sur des temps très courts (congélation immédiate) dans l'étude de la ficoline M, et des temps de trempage longs (nuit, jours) à des concentrations un peu plus faibles (50mM, 150mM) dans les études de la ficoline L et de gC1q.

F. Conclusions et perspectives.

Je ne reviendrai pas en détail sur les conclusions évoquées à la fin des chapitres C et D sur l'analyse structurale des complexes initiant les voies classique et lectine, respectivement. Il est clair que les recherches effectuées ces 10 dernières années ont apporté leur lot de surprises. Un résultat clair est néanmoins que ces complexes fonctionnent en se fixant avec avidité sur une surface 'cible' et que, si cette fixation entraîne un signal mécanique par déplacement des tiges de type collagène, alors elle pourra déclencher l'activation des protéases, et donc initier le démarrage de la cascade protéolytique.

En participant à divers congrès pour présenter ces résultats, j'ai pu progressivement découvrir d'autres aspects du système du complément, ce qui a nourri ma curiosité sur ce sujet. Au-delà du mécanisme perfectionné de la machinerie catalytique qui peut mener à la destruction et/ou à l'élimination efficace de bactéries, comme les streptocoques, la molécule C1q supporte plusieurs rôles intrigants. Parmi les aspects qui m'interpellent le plus, je citerai l'effet protecteur et régulateur au niveau de la réponse adaptative, qui apparaît clairement dans le risque associé aux déficits en C1q de développer des maladies autoimmunes. Le(s) mécanisme(s) associé(s) à ce risque n'est pas encore clairement identifié(s), mais les recherches sur ce sujet font évoluer la perception que l'on peut avoir du rôle de cette molécule. En effet, au-delà de la distinction dichotomique soi / non-soi, il semble que la molécule C1q assure un rôle important dans l'élimination efficace et tolérante d'éléments du soi altéré ou vieilli comme les cellules apoptotiques. La molécule C1q et les éléments de l'activation précoce comme C4b sont sans doute impliqués dans une voie particulière de phagocytose, rapide et accompagnée de signaux de tolérance. Certains pathogènes, comme le VIH, semblent avoir développé en partie une stratégie d'infiltration dans les cellules du système immunitaire qui utilise justement cette voie particulière. Comment qualifier alors la capacité de C1q de se fixer sur ses protéines externes gp41 et gp120 du VIH, qui semble porter deux effets contradictoires que sont la fonction originelle d'élimination et celle de facilitation de l'entrée du virus, liée aux mécanismes de subversion du système par le virus ? Parmi les autres propriétés qui m'intriguent, se trouve, en filigrane, un tropisme vers les organes lymphatiques et vers le cerveau, qui paraît corrélé à la fixation de C1q. Ce qui se traduirait par un effet facilitateur de la progression et pénétration dans le cerveau de certains éléments pathogènes (prion, VIH, ...). Ainsi a pu évoluer, avec le temps et grâce aux efforts de recherche d'une communauté, la perception que l'on peut avoir du rôle du système du complément et de ses interactions avec d'autres éléments de la défense innée, de la réponse adaptative, et même d'autres fonctions.

D'autres fonctions de la molécule C1q passent par son interaction avec des récepteurs cellulaires, mais les mécanismes précis de ces activités restent encore largement méconnus.

Quelques exemples seront cités ci-dessous :

- a) C1q contribue au chimiotactisme et à la maturation des cellules dendritiques (Csomor et al., 2007), induisant une réponse de type Th1 pour les lymphocytes voisins.
- b) C1q induit le chimiotactisme, renforce la phagocytose et l'élimination des microbes, et induit la production d'espèces réactives de l'oxygène par les neutrophiles
- c) C1q induit l'activation cellulaire des cellules endothéliales et des plaquettes et renforce leurs propriétés d'adhésion.
- d) C1q module la fonction des lymphocytes B et T...

Pour la molécule C1q, j'ai donc finalement plutôt l'impression d'avoir juste exploré la « surface émergée de l'iceberg ». Plusieurs projets s'inscrivent dans la continuité des travaux réalisés jusqu'ici, certains visant des parties « immergées » encore non explorées.

Les analyses structurales ont permis de montrer jusqu'à présent que le système du complément utilise une grande variété de mécanismes moléculaires pour contrôler l'activation des protéases, l'association des complexes ou les reconnaissances moléculaires (activatrices ou inhibitrices). Ce sont souvent les mêmes structures modulaires qui sont utilisées, comme les domaines fibrinogènes pour les différentes ficolines, ou les modules CCP, qui modulent l'activité des protéases C1r, C1s et MASPs, mais peuvent aussi être présents en dizaines de copies sur une même molécule (CR1). Le repliement global de chacun de ces modules est prévisible, mais leurs propriétés fines d'interaction et la manière dont ils s'assemblent dans l'espace ne peuvent pas être prédits si facilement. D'où la nécessité de réaliser de nouvelles analyses structurales expérimentales qui apporteront des informations sur les propriétés d'interaction. La cristallographie est une technique de choix pour observer précisément les interactions protéine-protéine ou protéine-ligand. Mais cette approche sous-entend quelques contraintes expérimentales qu'il est important de vérifier, faute de quoi le risque est de ne pas pouvoir aboutir :

- ◇ possibilité d'obtenir des protéines pures et très concentrées
- ◇ constantes d'affinité suffisantes entre les partenaires du complexe étudié
- ◇ si possible, que le complexe soit identifiable en solution.

J'évoquerai donc ci-dessous de manière très succincte quelques lignes directrices des projets en cours ou envisagés.

F.1. Propriétés de reconnaissance de C1q et des ficolines.

Avec l'aide de Virginie Garlatti, Lydie Martin et Anne Chouquet.

Collaborations actuelles : groupe de Gérard Arlaud (IBS/LEM), de Teizo Fujita (Japon), Berhane Ghebrehiwet (Etats-Unis).

Financement : ANR

Le code de reconnaissance utilisé par les protéines de l'immunité pour détecter les éléments à éliminer et déclencher la réponse appropriée est encore largement méconnu. Nous avons vu précédemment que, dans le système du complément, l'étape de reconnaissance conditionne à la fois les propriétés d'opsonisation et celles d'activation des voies classique et lectine. Nous avons bien progressé ces dernières années en déterminant les structures des domaines de reconnaissance de C1q et ceux des ficolines. Si l'on compare la structure globale des domaines trimériques de reconnaissance du complément (Figure F.1), on peut constater que les trimères de ficoline présentent une forme semi-ouverte, intermédiaire entre la forme très compacte de C1q et la forme très ouverte de la MBL. Seule la MBL a besoin d'une région supplémentaire (appelée 'neck') en amont du domaine de type lectine pour assembler le trimère.

Les distances entre les sites S1 des domaines globulaires de ficoline sont du même ordre de grandeur que celles observées entre les sites de la MBL, ce qui suggère que cela correspond à une distance assez fréquente entre les sucres terminaux des surfaces à reconnaître. Les sites additionnels découverts sur la ficoline L présentent par contre des propriétés géométriques différentes, et il est difficile d'envisager une fixation simultanée sur les trois sous-unités, jusqu'au site interne S2 et sans encombrement. L'écart entre les têtes globulaires modulera lui aussi la reconnaissance de motifs répétitifs à la surface des micro-organismes en fonction de leur distribution spatiale. Il serait sans doute intéressant de compléter les analyses cristallographiques par des études de microscopie sur des supports modèles ou réels permettant de visualiser la géométrie des motifs reconnus.

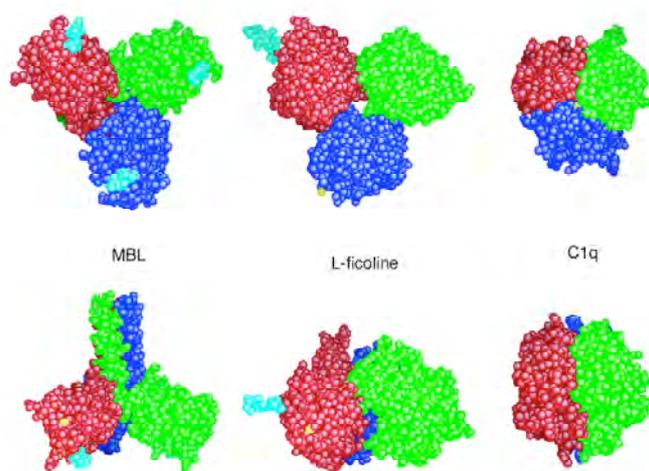


Figure F.1. Structures des têtes globulaires de reconnaissance de la MBL, L-ficoline et C1q complément, vue de dessus et de côté, à la même échelle. C1q a la forme la plus compacte. Les sites de fixation des sucres sur la MBL (bleu clair) sont du côté externe, à proximité du site de fixation de calcium (jaune).

Nous avons pu observer plusieurs exemples d'interactions dans le cas de l'étude des ficolines. Cette étude peut encore s'enrichir de l'analyse d'interactions avec d'autres ligands. Soit des ligands simples, comme les ligands acétylés pour la ficoline H ; soit des ligands plus complexes pour explorer les propriétés de reconnaissance de ligands étendus de la ficoline L ... Notre projet ANR inclut une recherche de nouveaux ligands par criblage systématique d'une banque de composés glycosidiques avec les ficolines (sous forme de molécules entières), en partenariat avec un consortium international de 'glycomique' fonctionnelle. Cependant, il faut garder en mémoire que les conditions expérimentales permettant de visualiser les interactions entre les ficolines et leurs ligands n'ont pas été très faciles à obtenir, une des causes étant vraisemblablement l'affinité modeste des ligands analysés pour les domaines globulaires isolés.

Nous souhaitons également continuer à explorer les propriétés de reconnaissance du domaine globulaire de C1q. Nous sommes confrontés sur ce projet au même problème d'affinité que celui évoqué ci-dessus avec les domaines isolés de ficolines, puisque C1q reconnaît une très grande variété de cibles en se fixant avec avidité sur des sites multiples: cibles disposées régulièrement et densément sur une surface (comme par exemple la phosphatidylsérine à la surface des cellules apoptotiques), ou cibles présentant une structure multimérique (IgM, Protéine C Réactive, SIGN-R1, ...) ou agrégée (réseau d'anticorps). Nous avons réalisé plusieurs séries d'essais de cristallisation sur des sujets qui nous intéressaient particulièrement, comme l'interaction avec le récepteur gC1qR, ou avec la protéine PrP. Dans les deux cas, les propriétés d'interaction ont été confirmées par BIAcore au niveau de la protéine C1q entière, mais nous n'avons pour l'instant aucun réel succès. D'autres cibles sont envisagées comme la MOG (Myelin Oligodendrocyte-Glycoprotein), pour laquelle un contrôle préalable des propriétés d'interaction avec le domaine globulaire de C1q devra être envisagé dans un premier temps. Nous essayons aussi de visualiser l'interaction de C1q avec de plus petits ligands, comme le N-acétyl-glucosamine ou les héparanes sulfates, pour lesquels il est possible d'utiliser la méthode alternative de trempage des cristaux en utilisant de fortes concentrations pour compenser la faible affinité du domaine globulaire isolé de C1q.

Les associations entre C1q et les protéines multimériques comme SIGN-R1, IgM ou les pentraxines CRP, SAP et PTX3 représentent des systèmes modèles très intéressants mais dont l'analyse structurale nécessitera d'utiliser d'autres techniques que la cristallographie.

F.2. Assemblage et activation des complexes C1 et ficolines/MASPs.

Aide technique : Anne Chouquet.

F.2.1. Valider le modèle d'activation du complexe C1

Collaboration: Gérard Arlaud, Véronique Bally, Thomas Lunardi (IBS/LEM).

Le modèle de C1 a été établi par association des structures des fragments ou modèles en respectant le modèle macroscopique établi préalablement. Le modèle ainsi construit est évidemment une version simplifiée de la molécule originale, présentant une symétrie exagérée de la position des bras. Pour améliorer notre compréhension de ce complexe, nous tentons de cristalliser directement la forme proenzyme du complexe C1. Des essais de mutagenèse dirigée sont réalisés en parallèle dans le groupe de Gérard Arlaud pour sonder par cette voie la validité du modèle proposé. En particulier, des mutations sont réalisées dans la boucle spécifiquement étendue et chargée négativement du module EGF de la molécule C1r pour essayer d'inhiber la formation du complexe C1.

F.2.2. Mieux comprendre l'assemblage des complexes ficolines/MASPs

Collaboration: Nicole Thielens, Gérard Arlaud (IBS/LEM).

Les derniers résultats obtenus sur le fragment CUB-EGF-CUB de MASP1/3 sont en cours d'analyses. Plusieurs acides aminés impliqués dans l'interaction avec la MBL et les différentes ficolines ont été mis en évidence par mutagenèse dirigée. De même, une lysine essentielle à l'interaction a été mise en évidence sur la MBL et les ficolines. Nous allons pouvoir analyser ces informations pour en déduire les modèles possibles d'interaction.

F.2.3. Visualiser les interactions moléculaires du domaine de type 'serpin' de C1-inhibiteur avec C1r, C1s, ...

Collaborations : Véronique Rossi (IBS/LEM), Christan Drouet (CHU, Grenoble)

Le dernier élément qui intervient dans l'activation du complexe C1 et sa dissociation est l'inhibiteur physiologique C1-inh (C1 inhibitor). Nous essayons de cristalliser un complexe formé par le domaine de type 'serpin' de cet inhibiteur et le fragment CCP2-SP de C1s. L'intérêt de ce projet est que C1-inh est le seul inhibiteur capable d'interagir avec C1r, C1s, les MASPs, et qu'il inhibe aussi des protéases du système de contact comme la kallikréine et le facteur XIIa, ainsi qu'une protéase de la cascade de la coagulation, le facteur XIa.

La structure de C1-inh vient d'être résolue très récemment par le groupe de Peter Zavodsky, dans une conformation inactive, dite 'latente' (Beinrohr et al., 2007). Cette structure présente une conformation particulière, non observée dans d'autres formes latentes de domaines de type 'serpin'. Cette structure ne permet pas de visualiser directement la conformation inhibitrice de C1-inh, et encore moins ses interactions avec les différentes protéases. Il s'agit là encore d'un projet difficile, les complexes protéases-'serpin' étant très difficiles à cristalliser, et la molécule C1-inh présentant une hétérogénéité de conformations due au phénomène de latence observée par nos collègues hongrois.

De nombreuses mutations ont été identifiées et analysées sur cette protéine, et plusieurs d'entre elles ont des répercussions pathologiques, étant associées à des angioedèmes héréditaires. Les positions des mutations ponctuelles observées seront situées sur la structure, afin d'aider à mieux percevoir les zones affectées. La connaissance de ces mutations et de leur impact physiologique pourra ainsi aider à mieux percevoir les régions fonctionnellement essentielles.

F.3. Une voie particulière de phagocytose des cellules apoptotiques ?

Collaboration : P. Frachet, IBS/LEM ;

J.P.Kleman, D. Housset, J.-B. Reiser, V. Garlatti, IBS/PSB

Il s'agit là d'un projet dont l'émergence est beaucoup plus récente, et qui tourne autour de la question de l'exploration des voies de reconnaissance et signalisation particulières impliquées dans la phagocytose des cellules apoptotiques. Je suis plus particulièrement concernée par les analyses structurales liées la reconnaissance moléculaire des cellules apoptotiques par C1q et les ficolines, et, dans un deuxième temps, à celles concernant les interactions moléculaires entre C1q ou les ficolines et les récepteurs impliqués, à la surface des phagocytes. J.P.Kleman, D. Housset, J.-B. Reiser sont concernés plus particulièrement par les études des interactions impliquées dans les cascades de signalisation associées.

Philippe Frachet et Héléna Païdassi (IBS/LEM) ont exploré les propriétés de reconnaissance des molécules de phosphatidylsérine par C1q, et nous avons collaboré pour essayer de visualiser l'interaction de reconnaissance réalisée au niveau des domaines globulaires de C1q. Le motif moléculaire reconnu est le groupement phosphosérine et nous avons pu proposer un site d'interaction (article en cours de révision). D'autres marqueurs moléculaires sont potentiellement impliqués dans ce processus de reconnaissance (CRP, IgM, calréticuline, ...), ce qui n'est pas contradictoire, vu le potentiel de liaisons multiples de la molécule C1q.

La question délicate est de savoir ensuite quel est le(s) récepteur(s) de C1q impliqué(s) dans le déclenchement de la phagocytose et de la signalisation associée, et d'identifier éventuellement à ce niveau des particularités liées à la reconnaissance de cellules apoptotiques. On peut également se demander s'il y a activation de C3, ou seulement de C4, pour faciliter le processus de phagocytose, en impliquant d'autres récepteurs comme CR3?

Pourquoi la question de l'identification de la nature et du rôle des récepteurs de C1q est-elle compliquée ? Parce qu'il n'existe pas un récepteur unique avec une action bien définie, facile à identifier, comme par exemple dans le cas des récepteurs de type Toll (Bohlsou et al., 2007). Deux ligands de C1q, d'origine intracellulaire, mais que l'on peut trouver exposés à la surface de différentes cellules, ont été proposés comme des récepteurs, même si cette appellation est controversée vue la nature des ligands (Ghebrehiwet et Peerschke, 2004) : il s'agit du récepteur des têtes globulaires (gC1qR ou p33) et du récepteur des tiges collagène cC1qR, plus connu sous le nom de calréticuline. Ces récepteurs n'étant pas des protéines membranaires, le problème est de trouver la molécule adaptatrice qui transmet le signal, laquelle pourrait varier en fonction du rôle étudié et des cellules impliquées. La molécule adaptatrice classiquement évoquée pour la calréticuline est la protéine CD91, mais CD59 pourrait également jouer ce rôle, au moins dans certaines conditions (Ghiran et al., 2003). Comme nous l'avons vu dans l'introduction, deux autres récepteurs sont à prendre en considération: l'intégrine $\alpha 2\beta 1$ et le récepteur CR1. Et d'autres sont peut-être encore à découvrir, comme par exemple un homologue humain de SIGN-R1..?

Le dernier élément, qui finit de compliquer le tableau, est que le contact établi sur les phagocytes par la molécule C1q (ou les cellules apoptotiques) semble déclencher la montée en surface de molécules très importantes pour la phagocytose, colocalisées dans les 'radeaux' membranaires. On peut citer, dans cette catégorie, le transporteur ABCA7 et la molécule CD91 (Jehle et al, 2006), et peut-être aussi, dans d'autres circonstances, le récepteur CD93. Se greffe donc le rôle de stimulus de C1q dans la présentation de récepteurs en surface, qui n'implique pas forcément de contact direct avec les récepteurs ainsi mobilisés...

Dans le cas particulier de la phagocytose des cellules apoptotiques par des macrophages, une première piste à explorer est celle de l'interaction entre les tiges de C1q ou des collectines et la calréticuline. Cependant la calréticuline n'est pas une molécule facile à

cristalliser (de nombreux essais ont déjà été réalisés dans d'autres équipes), ni à manipuler. C'est une protéine multifonctionnelle et flexible, et donc il faudrait s'intéresser plutôt dans un premier temps au fragment responsable de la fixation de C1q, le domaine 'S'. Du côté des tiges collagène, on peut envisager de faire synthétiser un fragment qui inclut la séquence la plus conservée parmi les différentes protéines : GE(K/Q/R)GEP. Pour l'étude de la signalisation, la protéine adaptatrice de CD91 semble être GULP, et donc nous recherchons ses partenaires.

Une autre piste qui pourrait être intéressante à approfondir est celle de la molécule CR1. En effet, C1q, C4 et CR1 font partie des gènes dont la déficience est fortement corrélée à une prédisposition au développement de lupus systémique érythémateux (SLE), pour laquelle une déficience dans l'élimination correcte des cellules apoptotiques a été supposée. De plus, CR1 porte à la fois un site de liaison à C4b et à C1q, la fixation combinée sur les deux sites étant additive (Tas et al., 1999). Cette hypothèse d'une implication combinée de ces 3 gènes a été réfutée initialement en utilisant des souris déficientes en CR1/2, lesquelles ne développent pas ce genre de symptôme (Chen et al., 2000). Le problème est que la fonction portée par le gène CR1 chez l'homme n'est pas identique à celle du gène CR1/2 (1 seul gène) chez la souris : la fonction de prise en charge des complexes immuns étant effectuée par le facteur H chez la souris, à la place du CR1 chez l'homme (Alexander et Quigg, 2007). Le site de fixation de C1q n'a pas été identifié précisément. Celui-ci semble localisé dans le domaine D, le plus proche de la membrane cytoplasmique, sur les modules CCP 22 à 28. Un nouvel anticorps reconnaissant des épitopes conservés des modules CCP 28-29, présents dans toutes les formes de CR1, vient d'être décrit (Chen et al., 2007), et il serait intéressant de tester s'il a une influence sur la liaison de C1q. De même, il est possible de préciser le site de liaison pour C1q dans le domaine D par production de sous-domaines recombinants plus ou moins longs en N-terminal. Une fois identifié, nous pouvons envisager d'analyser la structure du plus petit domaine recombinant contenant le site de fixation de C1q, et de l'utiliser dans des expériences de compétition. Enfin un intérêt du récepteur CR1 est qu'il dispose d'une partie cytoplasmique, ce qui permet d'identifier des partenaires fonctionnels dans la signalisation, comme cela a été entrepris pour CR2 (Tessier et al., 2007).

F.4. Pathogènes et récepteurs du complément.

Ce projet complète le point précédent. L'idée est de s'intéresser aux mêmes molécules (calréticuline, CR1, gC1qR...), mais cette fois sous l'angle des mécanismes utilisés par les pathogènes pour contrer et même utiliser à leur profit le système immunitaire. Un des intérêts est de faire apparaître des projets plus orientés vers des applications thérapeutiques potentielles, pour lesquels nous pourrions sans doute plus facilement rédiger des demandes de financement. L'autre intérêt est de tirer partie du fait que les pathogènes ont évolué en parallèle de notre système de défense, et que nous pouvons apprendre par ce biais les points forts et les fragilités de notre immunité innée. Ces projets s'intégreront parfaitement bien dans le nouvel axe de recherche de l'IBS intitulé 'Immunité et interactions hôte-pathogènes'.

a) CR1 et malaria :

CR1 présente un site de fixation pour une adhésine du *Plasmodium*, la protéine de surface PfEMP1 (*Plasmodium falciparum* erythrocyte membrane protein 1). Celle-ci semble représenter un facteur de virulence important, favorisant la formation de rosettes de globules rouges et l'adhésion aux micro-vaisseaux. Un site de fixation pour cette adhésine a été localisé dans les CCP 15-17 du CR1 (Krych-Goldberg et al., 2002). Un polymorphisme très important du gène CR1, touchant la région D (CCP25), semble protéger des formes graves de la maladie, comme la malaria cérébrale et les anémies sévères (Cockburn et al., 2004).

Réciproquement, de très nombreuses formes existent pour la protéine PfEMP1, et sont associées à des degrés de pathologie très différents (Kaestli et al., 2006).

b) Calréticuline du trypanosome (maladie de Chagas) et C1q humain.

La calréticuline est une protéine très conservée à travers les espèces, et on peut la retrouver exprimée à la surface de divers microbes et parasites, dont le plasmodium (*Plasmodium falciparum*), mais aussi le trypanosome. Sachant que les cellules apoptotiques exposent la calréticuline à la surface, et que les cellules cancéreuses modulent aussi l'expression de cette protéine, il y a un véritable intérêt à regarder de plus près le rôle de cette protéine. C'est ce qui a été fait dans le cas du trypanosome (Ferreira et al., 2005). La calréticuline semble porter dans ce cas un double rôle, celui de protection contre l'activation du complément par la voie classique, et aussi celui d'agent inhibiteur de l'angiogénèse, qui constitue une protection de l'hôte contre la progression métastatique des cancers.

Le fragment recombinant TcCRT R (aa 136–281) de la calréticuline de *Trypanosoma cruzi* (code d'accèsion Entrez AF162779), relativement stable, permet de reproduire la propriété de fixation et d'inhibition de l'activation de la voie classique par la calréticuline (Ferreira et al., 2004). Il s'agit donc là d'un candidat intéressant pour nos études.

c) Protéine du virus de l'hépatite C et récepteurs des têtes globulaires de C1q.

L'hépatite C représente un enjeu important de santé publique puisqu'on estime à environ 170 millions le nombre de porteurs dans le monde. Nous pouvons envisager l'étude du complexe formé entre la protéine gC1qR et une protéine exprimée par le virus de l'hépatite C, appelée 'hepatitis C virus core protein'. Cette protéine, la première à être exprimée et à circuler dans le sang des personnes infectées, inhibe la réponse proliférative des lymphocytes T en se fixant de façon spécifique sur le gC1qR à la surface des cellules T, ce qui doit contribuer à aider à la survie du virus durant les phases aiguës d'infection (Yao et al., 2001). Un effet clair de cette fixation a été noté concernant l'infection chronique ou non dans le cas des chimpanzés, qui représentent un système très intéressant, car la moitié d'entre eux résiste à l'infection chronique (Yao et al., 2006). De plus, ces mêmes auteurs ont montré une répression de l'expression de l'interleukine 12 (signal inflammatoire) par le pontage des récepteurs gC1qR, par inhibition de la voie de signalisation Lck et Akt (Yao et al., 2004 ; Waggoner et al., 2005). Cette protéine de l'hépatite C est étudiée à l'IBCP à Lyon. Il serait intéressant de réussir à localiser la fonction d'interaction avec gC1qR dans cette protéine et voir si ce système serait plus favorable pour une étude cristallographique.

Pour conclure, j'espère pouvoir renforcer assez rapidement mon groupe avec un chercheur permanent, de manière à être en mesure de mettre en œuvre une partie importante des projets évoqués ci-dessus, et d'établir de nouvelles collaborations. Certains de ces projets seront repris et développés dans de nouvelles demandes de financement.

REFERENCES

- Alexander J.J., Quigg R.J. The simple design of complement factor H: Looks can be deceiving. (2007) *Mol. Immunol.* **44**, 123-132.
- Aoyagi Y., Adderson E.E., Min J.G., Matsushita M., Fujita T., Takahashi S., Okuwaki Y., Bohnsack J.F. (2005) Role of L-ficolin/mannose-binding lectin-associated serine protease complexes in the opsonophagocytosis of type III group B streptococci. *J Immunol.* **174**, 418-425.
- Arlaud G.J., Colomb M.G., Gagnon J. (1987) A functional model of the human C1 complex : Emergence of a functional model. *Immunology Today*, **8**, 106-111
- Banki Z., Wilflingseder D., Ammann C.G., Pruenster M., Mullauer B., Hollander K., Meyer M., Sprinzl G.M., van Lunzen J., Stellbrink H.J., Dierich M.P., Stoiber H. (2006) Factor I-mediated processing of complement fragments on HIV immune complexes targets HIV to CR2-expressing B cells and facilitates B cell-mediated transmission of opsonized HIV to T cells. *J. Immunol.*, **177**, 3469-3476.
- Bao L., Quigg R.J. (2007) Complement in lupus nephritis: the good, the bad, and the unknown. *Semin. Nephrol.*, **27**, 69-80
- Behrens E.M., Sriram U., Shivers D.K., Gallucci M., Ma Z., Finkel T.H., Gallucci S. (2007) Complement Receptor 3 Ligation of Dendritic Cells Suppresses Their Stimulatory Capacity. *J. Immunol.*, **178**, 6268-6279.
- Beinrohr L., Harmat V., Dobo J., Lorincz Z., Gal P., Zavodszky P. (2007) C1-inhibitor serpin domain structure reveals the likely mechanism of heparin potentiation and conformational disease. *J. Biol. Chem.* Sous presse, version web accessible.
- Bohlsón S.S., Fraser D.A., Tenner A.J. Complement proteins C1q and MBL are pattern recognition molecules that signal immediate and long-term protective immune functions. *Mol. Immunol.*, **44**, 33-43.
- Bureeva S., Andia-Pravdivy J., Kaplun A. (2005) Drug design using the example of the complement system inhibitors' development. *Drug Discov. Today*, **10**, 1535-1542.
- Burton D.R., Boyd J., Brampton A.D., Easterbrook-Smith S.B., Emanuel E.J., Novotny J., Rademacher T.W., van Schravendijk M.R. and Dwek R.A. (1980) The C1q receptor site on immunoglobulin G. *Nature* , **288**, 338-344.
- Chen C.H., Ghiran I., Beurskens F.J., Weaver G., Vincent J.A., Nicholson-Weller A., Klickstein L.B. (2007) Antibody CR1-2B11 recognizes a non-polymorphic epitope of human CR1 (CD35). *Clin. Exp. Immunol.*, **148**, 546-54.

- Chen Z., Koralov S.B., Kelsoe G. Complement C4 inhibits systemic autoimmunity through a mechanism independent of complement receptors CR1 and CR2. *J. Exp. Med.*, **192**, 1339-1352.
- Cockburn I.A., Mackinnon M.J., O'Donnell A., Allen S.J., Moulds J.M., Baisor M., Bockarie M., Reeder J.C., Rowe J.A. (2004) A human complement receptor 1 polymorphism that reduces Plasmodium falciparum rosetting confers protection against severe malaria. *Proc. Natl. Acad. Sci U S A.*, **101**, 272-277.
- Cooper N.R. (1985) The classical complement pathway: activation and regulation of the first complement component. *Adv. Immunol.*, **37**, 151-216.
- Crocker P.R., Paulson J.C., Varki A. (2007) Siglecs and their roles in the immune system. *Nat. Rev. Immunol.*, **7**, 255-266.
- Cseh S., Vera L., Matsushita M., Fujita T., Arlaud G.J., Thielens NM. (2002) Characterization of the interaction between L-ficolin/p35 and mannan-binding lectin-associated serine proteases-1 and -2. *J. Immunol.*, **169**, 5735-573.
- Csomor E., Bajtay Z., Sandor N., Kristof K., Arlaud G.J., Thiel S., Erdei A. (2007) Complement protein C1q induces maturation of human dendritic cells. *Mol. Immunol.* **44**, 3389-3397
- Dempsey P.W., Allison M.E.D., Srinivas A., Goodnow C.C. and Fearon D.T. (1996) C3d of complement as a molecular adjuvant: Bridging innate and acquired immunity. *Science* **271**, 348-350.
- Dall'Acqua W.F., Cook K.E., Damschroder M.M., Woods R.M., Wu H. (2006) Modulation of the effector functions of a human IgG1 through engineering of its hinge region. *J. Immunol.*, **177**, 1129-1138.
- Edelson B.T., Li Z., Pappan L.K., Zutter M.M. (2004) Mast cell-mediated inflammatory responses require the alpha 2 beta 1 integrin. *Blood*, **103**, 2214-2220
- Edelson B.T., Stricker T.P., Li Z., Dickeson S.K., Shepherd V.L., Santoro S.A., Zutter M.M. (2006) Novel collectin/C1q receptor mediates mast cell activation and innate immunity. *Blood*, **107**, 43-50.
- Endo Y., Nakazawa N., Liu Y., Iwaki D., Takahashi M., Fujita T., Nakata M., Matsushita M. (2005) Carbohydrate-binding specificities of mouse ficolin A, a splicing variant of ficolin A and ficolin B and their complex formation with MASP-2 and sMAP. *Immunogenetics*. **57**, 837-844.
- Endo Y., Takahashi M., Fujita T. (2006) Lectin complement system and pattern recognition. *Immunobiology*, **211**, 283-293.

- Feinberg H., Uitdehaag J. C. M., Davies J. M., Wallis R., Drickamer K., and Weiss W. I. (2003) Crystal structure of the CUB1-EGF-CUB2 region of mannan-binding protein associated serine protease-2. *EMBO J.* **22**, 2348-2359
- Ferreira V., Valck C., Sanchez G., Gingras A., Tzima S., Molina M.C., Sim R., Schwaeble W., Ferreira A. (2004) The classical activation pathway of the human complement system is specifically inhibited by calreticulin from *Trypanosoma cruzi*. *J. Immunol.*, **172**, 3042–3050.
- Ferreira V., Molina M.C., Schwaeble W., Lemus D., Ferreira A. (2005) Does *Trypanosoma cruzi* calreticulin modulate the complement system and angiogenesis? *Trends Parasitol.*, **21**, 169-74
- Fishelson Z., Donin N., Zell S., Schultz S., Kirschfink M. (2004) Obstacles to cancer immunotherapy : expression of membrane complement regulatory proteins in tumors. *Mol. Immunol.* **41**, 109-123.
- Fraser D.A., Arora M., Bohlsion S.S., Lozano E., Tenner A.J. (2007) Generation of inhibitory NFkappaB complexes and phosphorylated cAMP response element-binding protein correlates with the anti-inflammatory activity of complement protein C1q in human monocytes. *J Biol Chem.*, **282**, 7360-7367.
- Gasque P. (2004) Complement : a unique innate immune sensor for danger signals. *Mol. Immunol.*, **41**, 1089-1098.
- Ghebrehiwet B. and Peerschke E.I.B (2004) cC1q-R (calreticulin) and gC1qR/p33 : ubiquitously expressed multi-ligand binding cellular proteins involved in inflammation and infection. *Mol. Immunol.* **41**, 173-183.
- Ghiran I., Klickstein L.B., Nicholson-Weller A. (2003) Calreticulin is at the surface of circulating neutrophils and uses CD59 as an adaptor molecule.. *J. Biol. Chem.*, **278**, 21024-21031.
- Guddat L.W., Herron J.N. and Edmunson A.B. (1993) Three-dimensional structure of a human immunoglobulin with a hinge deletion. *Proc. Natl. Acad. Sci. USA* **90**, 4271-4275.
- Hezareh M., Hessel A.J., Jensen R.C., van de Winkel J.G.J. and Parren P.W.H.I. (2001) Effector function activities of a panel of mutants of a broadly neutralizing antibody against human immunodeficiency virus type 1. *J. Virol.* **75**, 12161-12168.
- Holland M.C.H., Lambris J.D. Complement (2005) Encyclopedia of Life Sciences, Nature Publishing Group, www.els.net.

- Idusogie E.E., Presta L.G., Gazzano-Santoro H., Totpal K., Wong P.Y., Ultsch M., Meng Y.G., Mulkerrin M.G. (2000) Mapping of the C1q binding site on rituxan, a chimeric antibody with a human IgG1 Fc. *J. Immunol.* **164**, 4178-4184.
- Idusogie E.E., Wong P.Y., Presta L.G., Gazzano-Santoro H., Totpal K., Ultsch M. and Mulkerrin M.G. (2001) Engineered antibodies with increased activity to recruit complement. *J. Immunol.* **166**, 2571-2575.
- Illy C., Thielens N. M., and Arlaud G. J. (1993) Chemical characterization and location of ionic interactions involved in the assembly of the C1 complex of human complement. *J. Prot. Chem.* **12**, 771-781.
- Jegou J.F., Chan P., Schouft M.T., Griffiths M.R., Neal J.W., Gasque P., Vaudry H., Fontaine M. (2007) C3d binding to the myelin oligodendrocyte glycoprotein results in an exacerbated experimental autoimmune encephalomyelitis. *J. Immunol.* **178**, 3323-3331.
- Jehle A.W., Gardai S.J., Li S., Linsel-Nitschke P., Morimoto K., Janssen W.J., Vandivier R.W., Wang N., Greenberg S., Dale B.M., Qin C., Henson P.M., Tall A.R. (2006) ATP-binding cassette transporter A7 enhances phagocytosis of apoptotic cells and associated ERK signaling in macrophages. *J. Cell. Biol.*, **174**, 547-456.
- Johnson S., Michalak M., Opas M., Eggleton P. (2001) The ins and outs of calreticulin : from the ER lumen to the extracellular space. *Trends Cell Biol.* **11**, 122-129.
- Kaestli M., Cockburn I.A., Cortes A., Baea K., Rowe J.A., Beck H.P. (2006) Virulence of malaria is associated with differential expression of Plasmodium falciparum var gene subgroups in a case-control study. *J Infect. Dis.*, **193**, 1567-1574.
- Kairies N., Beisel H.G., Fuentes-Prior P., Tsuda R., Muta T., Iwanaga S., Bode W., Huber R., Kawabata S.I. (2001) The 2.0-Å crystal structure of tachylectin 5A provides evidence for the common origin of the innate immunity and the blood coagulation systems. *Proc Natl Acad Sci USA* **98**: 13519-13524
- Kang Y.S., Do Y., Lee H.K., Park S.H., Cheong C., Lynch R.M., Loeffler J.M., Steinman R.M., Park C.G. (2006) A dominant complement fixation pathway for pneumococcal polysaccharides initiated by SIGN-R1 interacting with C1q. *Cell*, **125**, 47-58.
- Kilchherr E., Hofmann H., Steigemann W. and Engel J. (1985) Structural model of the collagen-like region of C1q comprising the kink region and the fibre-like packing of the six triple helices. *J. Mol. Biol.* **186**, 403-415.
- Knobel H.R., Villiger W., 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., Reid K.B., Kishore U. (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.* **172**, 4351-4358
- Krarrup A., Thiel S., Hansen A., Fujita T., Jensenius J.C. (2004) L-ficolin Is a Pattern Recognition Molecule Specific for Acetyl Groups., *J. Biol. Chem.*, **279**, 47513-47519
- Krych-Goldberg M., Moulds J.M., Atkinson J.P. Human complement receptor type 1 (CR1) binds to a major malarial adhesin. *Trends Mol. Med.* **8**, 531-537.
- Lacroix M., Ebel C., Kardos J., Dobo J., Gal P., Zavodszky P., Arlaud G.J. and Thielens N.M. (2001) Assembly and enzymatic properties of the catalytic domain of human complement protease C1r. *J. Biol. Chem.*, **276**, 36233-36240
- Lim J., Wiedemann A., Tzircotis G., Monkley S.J., Critchley D.R., Caron E. (2007) An essential role for talin during alpha(M)beta(2)-mediated phagocytosis. *Mol. Biol. Cell*, **18**, 976-985.
- Lynch N.J., Roscher S., Hartung T., Morath S., Matsushita M., Maennel D.N., Kuraya M., Fujita T., Schwaeble W.J. (2004) L-ficolin specifically binds to lipoteichoic acid, a cell wall constituent of Gram-positive bacteria, and activates the lectin pathway of complement. *J. Immunol.*, **172**, 1198-1202.
- Ma Y.G., Cho M.Y., Zhao M., Park J.W., Matsushita M., Fujita T., Lee B.L. (2004) Human mannose-binding lectin and L-ficolin function as specific pattern recognition proteins in the lectin activation pathway of complement. *J. Biol. Chem.*, **279**, 25307-25312.
- Marqués G., Anton L.C., Barrio E., Sanchez A., Ruiz A., 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.
- Mastellos D, Andronis C, Persidis A, Lambris J.D. (2005) Novel biological networks modulated by complement. *Clin. Immunol.*, **115**, 225-235.
- 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.
- Morgan B.P., Gasque P. (1997) Extrahepatic complement biosynthesis: where, when and why? *Clin Exp Immunol.*, **107**, 1-7.
- Morris R.J., Perrakis A., Lamzin V.S. (2003) ARP/wARP and automatic interpretation of protein electron density maps. *Methods Enzymol.*, **374**, 229-244.

- Nahid A.M., Sugii S. (2006) Binding of porcine ficolin-alpha to lipopolysaccharides from Gram-negative bacteria and lipoteichoic acids from Gram-positive bacteria. *Dev. Comp. Immunol.*, **30**, 335-343.
- Nauta A.J., Castellano G., Xu W., Woltman A.M., Borrias M.C., Daha M.R., van Kooten C., Roos A. (2004) Opsonization with C1q and Mannose-Binding Lectin targets apoptotic cells to dendritic cells. *J Immunol.*, **173**, 3044-3050.
- Peiser L., Gordon S. Phagocytosis : Enhancement, (2001) Encyclopedia of Life Sciences, Nature Publishing Group, www.els.net.
- Rainey J.K. and Goh M.C. (2002) A statistically derived parameterization for the collagen triple-helix. *Protein Sci.* **11**, 2748-2754.
- Reid K. B. M. (1979) Complete amino acid sequences of the three collagen-like regions present in subcomponent C1q of the first component of human complement. *Biochem. J.*, **179**, 367-371
- Riley-Vargas R.C., Lanzendorf S., Atkinson J.P. (2005) Targeted and restricted complement activation on acrosome-reacted spermatozoa. *J. Clin. Invest.* , **115**, 1241-1249
- Rossi V., Bally I., Thielens N.M., Esser A.F., 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.
- Rossi V., Cseh S., Bally I., Thielens N.M., Jensenius J.C., Arlaud G.J. (2001) Substrate specificities of recombinant mannan-binding lectin-associated serine proteases-1 and -2., *J. Biol. Chem.*, **276**, 40880-40887.
- Rubinstein D.B., Stortchevoi A., Boosalis M., Ashfaq R., Ghebrehiwet B., Peerschke E.I., Calvo F., Guillaume T. (2004) Receptor for the globular heads of C1q (gC1q-R, p33, hyaluronan-binding protein) is preferentially expressed by adenocarcinoma cells. *Int J Cancer.*, **110**, 741-50.
- Runza V.L., Hehlhans T., Echtenacher B., Zahringer U., Schwaeble W.J., Mannel D.N. (2006) Localization of the mouse defense lectin ficolin B in lysosomes of activated macrophages. *J. Endotoxin Res.* , **12**, 120-126.
- Sahin-Toth M., Kukor Z., Nemoda Z. (2006) Human cationic trypsinogen is sulfated on Tyr154. *FEBS J.*, **273**, 5044-5050.

- Saphire E.O., Parren P.W., Pantophlet R., Zwick M.B., Morris G.M., Rudd P.M., Dwek R.A., Stanfield R.L., Burton D.R. and Wilson I.A. (2001) Crystal structure of a neutralizing human IgG against HIV-1: a template for vaccine design. *Science*, **293**, 1155-1159.
- Stengaard-Pedersen K., Thiel S., Gadjeva M., Moller-Kristensen M., Sorensen R., Jensen L. T., Sjöholm A. G., Fugger L., and Jensenius J. C. (2003) Inherited deficiency of mannan-binding lectin-associated serine protease 2. *N. Engl. J. Med.*, **349**, 554-560
- Stoiber H., Pruenster M., Ammann C.G. and Dierich M. (2005) . Complement-opsonized HIV : the free rider on its way to infection. *Mol. Immunol.*, **42**, 153-160.
- Subasinghe N.L., Ali F., Illig C.R., Jonathan Rudolph M., Klein S., Khalil E., Soll R.M., Bone R.F., Spurlino J.C., DesJarlais R.L., Crysler C.S., Cummings M.D., Morris P.E. Jr, Kilpatrick J.M., Sudhakara Babu Y. (2004) A novel series of potent and selective small molecule inhibitors of the complement component C1s. *Bioorg. Med. Chem. Lett.*, **14**, 3043-3047.
- Subasinghe N.L., Travins J.M., Ali F., Huang H., Ballentine S.K., Marugán J.J., Khalil E., Hufnagel H.R., Bone R.F., DesJarlais R.L., Crysler C.S., Ninan N., Cummings M.D., Molloy C.J. and Tomczuk B.E. (2006) A novel series of arylsulfonylthiophene-2-carboxamidine inhibitors of the complement component C1s. *Bioorg. Med. Chem. Lett.*, **16**, 2200-2204.
- Suber F., Carroll M.C., Moore F.D. (2007) Innate response to self-antigen significantly exacerbates burn wound depth. *Proc. Natl. Acad. Sci. U S A*, **104**, 3973-3977.
- Tas S.W., Klickstein L.B., Barbashov S.F., Nicholson-Weller A. (1999) C1q and C4b bind simultaneously to CR1 and additively support erythrocyte adhesion. *J. Immunol.*, **163**, 5056-5063.
- Teillet F., Lacroix M., Thiel S., Weilguny D., Agger T., Arlaud G.J., Thielens N.M. (2007) Identification of the site of human mannan-binding lectin involved in the interaction with its partner serine proteases: the essential role of lys55. *J. Immunol.* **178**, 5710-5716.
- Tessier J., Cuvillier A., Glaudet F., Khamlichi A.A. (2007) Internalization and molecular interactions of human CD21 receptor. *Mol. Immunol.* **44**, 2415-2425.
- Urich E., Gutcher I., Prinz M., Becher B. (2006) Autoantibody-mediated demyelination depends on complement activation but not activatory Fc-receptors. *Proc. Natl. Acad. Sci. U S A*, **103**, 18697-18702.
- Vuillard L., Rabilloud T., Leberman R., Berthet-Colominas C., Cusack S. (1994). A new additive for protein crystallization. *FEBS Lett.*, **353**, 294-296.

- Vuillard L., Marret N. and Rabilloud T. (1995). Enhancing protein solubilization with non-detergent sulfobetaines. *Electrophoresis*, **16**, 295–297.
- Waggoner S.N., Cruise M.W., Kassel R., Hahn Y.S. (2005) gC1q receptor ligation selectively down-regulates human IL-12 production through activation of the phosphoinositide 3-kinase pathway. *J. Immunol.*, **175**, 4706-4714.
- Weis T.J., Fearon D.T., 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.
- Weis W.I., Drickamer K., Hendrickson W.A. (1992) Structure of a C-type mannose-binding protein complexed with an oligosaccharide. *Nature*, **360**, 127-134.
- Wicher K.B., Fries E. (2004) Prohaptoglobin is proteolytically cleaved in the endoplasmic reticulum by the complement C1r-like protein. *Proc Natl Acad Sci U S A.*, **101**, 14390-14395
- Wiedemann A., Patel J.C., Lim J., Tsun A., van Kooyk Y., Caron E. (2006) Two distinct cytoplasmic regions of the beta2 integrin chain regulate RhoA function during phagocytosis. *J Cell Biol.*, **172**, 1069-1079.
- Yao Z.Q., Nguyen D.T., Hiotellis A.I., Hahn Y.S. (2001) Hepatitis C virus core protein inhibits human T lymphocyte responses by a complement-dependent regulatory pathway. *J. Immunol.*, **167**, 5264-5272.
- Yao Z.Q., Eisen-Vandervelde A., Waggoner S.N., Cale E.M., Hahn Y.S. (2004) Direct binding of hepatitis C virus core to gC1qR on CD4+ and CD8+ T cells leads to impaired activation of Lck and Akt. *J Virol.*, **78**, 6409-6419.
- Yao Z.Q., Shata M.T., Tricoche N., Shan M.M., Brotman B., Pfahler W., Hahn Y.S., Prince A.M. (2006) gC1qR expression in chimpanzees with resolved and chronic infection: potential role of HCV core/gC1qR-mediated T cell suppression in the outcome of HCV infection. *Virology*, **346**, 324-337.
- Zhang J., Wright I.W., Bernlohr D.A., Cushman S.W., and Chen X. (2007) Alterations of the classic pathway of complement in adipose tissue of obesity and insulin resistance. *Am. J. Physiol. Endocrinol. Metab.*, **292**, E1433-E1440
- Zhou W., Peng Q., Li K., Sacks S.H. (2007) Role of dendritic cell synthesis of complement in the allospecific T cell response. *Mol. Immunol.*, **44**, 57-63.
- Ziller F., Macor P., Bulla R., Sblattero D., Marzari R., Tedesco F. (2005) Controlling complement resistance in cancer by using human monoclonal antibodies that neutralize complement-regulatory proteins CD55 and CD59. *Eur. J. Immunol.*, **35**, 2175-2183.