Mass Spectrometry Analysis of the Oligomeric C1q Protein Reveals the B Chain as the Target of Trypsin Cleavage and Interaction with Fucoidan†

Bérangère Tissot,‡ Florence Gonnet,§ Alban Iborra,§ Christian Berthou,‡ Nicole Thielens,|| Gérard J. Arlaud,|| and Régis Daniel*,§

Laboratoire Analyse et Environnement, UMR 8587 CNRS, Université d'Evry-Val-d'Essonne, Boulevard F. Mitterrand, 91025 Evry Cedex, France, Laboratoire de Thérapie Cellulaire, CHU Morvan, 5 Avenue Foch, 29200 Brest, France, and Laboratoire d'Enzymologie Moléculaire, Institut de Biologie Structurale Jean-Pierre Ebel, CEA-CNRS-Université Joseph Fourier, 41 Rue Jules Horowitz, 38027 Grenoble Cedex 1, France

Received October 14, 2004; Revised Manuscript Received November 25, 2004

ABSTRACT: C1q is a subunit of the C1 complex that triggers activation of the complement classical pathway through recognition and binding of immune complexes. C1q also binds to nonimmune ligands such as the sulfated polysaccharide fucoidan, a potent anticomplementary agent. C1q was submitted for the first time to mass spectrometry analysis, yielding insights into its assembly and its interaction with fucoidan. The MALDI-TOF mass spectrometry technique on membrane allowed partial preservation of noncovalent interactions, allowing precise analysis of its substructure and estimation of the C1q molecular weight at 459520−461883, with an average mass of 460793 g·mol^{-1}. The disulfide-linked A-B and C-C dimers as well as the noncovalent structural unit (A-B:C)-(C:B-A) were detected, providing experimental support to the C1q model based on covalent and noncovalent associations of six heterotrimers. Trypsin treatment of native C1q led to proteolysis of the B chain only, at a single cleavage site (Arg^{109}) located in the globular region. Unlike DNA, fucoidan protected C1q from trypsin cleavage, indicating that this polysaccharide binds to the B moiety of the globular head. Given the involvement of the C1q globular heads in the recognition of IgG, this interaction may account for the observed anticomplementary activity of fucoidan.

Algal fucoidan exhibits a potent inhibitory activity against the human complement system. This property is of major interest given the involvement of complement in numerous pathological processes and the strong demand for efficient anticomplementary molecules (*1*). Unraveling the mechanism of action of a bioactive polysaccharide such as fucoidan requires a fine knowledge of its carbohydrate structure and of its binding to the target protein. Much effort has been made during the past years to elucidate the structure of fucoidan (*2*). The widely studied fucoidan from the brown alga *Ascophyllum nodosum* has a →4)-α-L-Fuc*p*-(1→3)-α-L-Fuc*p*-(1→4)-α-L-Fuc*p*-(1→3)-α-L-Fuc*p*(1→ backbone with a sulfate group mainly on the C2 position of the fucosyl units (*3*, *4*). Much less is known about its mechanism of interaction with its ligands. We have recently reported that *A. nodosum* fucoidan inhibits the first steps of the complement cascade through specific interactions with target proteins, particularly C1q (*5*).

C1q is a multimeric glycoprotein comprising 18 chains of three different types, A, B, and C (217−226 residues). These chains form three C-C and six A-B disulfide-bonded dimers, which assemble into six identical heterotrimeric (A-B:C) subunits through noncovalent association (A-B:C)-(C:B-A). This hexameric structure appears in electron microscopy as a bouquet of six flowers and exhibits unique structural features. It consists of a short N-terminal region (3−9 residues) containing the interchain (A-B and C-C) disulfide bonds, prolonged by a triple-helical collagen-like region (CLR,[1] about 81 residues) which forms a "stalk" and then diverges into six individuals "stems", each terminating in a C-terminal heterotrimeric globular "head" (GR, about 135 residues) (*6*, *7*) (Figure 1A).

C1q mediates binding of the C1 complex to antigen−antibody complexes and thereby triggers complement activation (*7*). Interaction with immune complexes involves the globular region of C1q (*8*, *9*) as recently detailed in a structural model derived from the X-ray structure of the C1q globular head (*10*). This region also mediates direct binding of apoptotic cells, giving C1q a major role in the clearance of aberrant structures from self (*11*). The collagen-like region, on the other hand, is also endowed with binding properties. First, it binds the catalytic C1s-C1r-C1r-C1s tetramer through the "stem" region (Figure 1B). Antibody-dependent activa-

† This work was supported by the Génopole, the CNRS, and the Université Evry Val d'Essonne. R.D. is indebted to the Génopole for financial support through the Action Thématique Incitative de Génopole.

* Correspondence should be addressed to this author. Tel: +33-1-69-47-7641. Fax: +33-1-69-47-7655. E-mail: regis.daniel@chimie.univ-evry.fr.

‡ Laboratoire de Thérapie Cellulaire, CHU Morvant.

§ Laboratoire Analyse et Environnement, UMR 8587 CNRS, Université d'Evry-Val-d'Essonne.

|| Laboratoire d'Enzymologie Moléculaire, Institut de Biologie Structurale Jean-Pierre Ebel, CEA-CNRS-Université Joseph Fourier.

[1] Abbreviations: CLR, collagen-like region; GR, globular head region; DHB, 2,5-dihydroxybenzoic acid; HCCA, α-cyano-4-hydroxycinnamic acid; SA, sinapinic acid; au, arbitrary units.

10.1021/bi047802h CCC: $30.25
Published on Web 02/15/2005

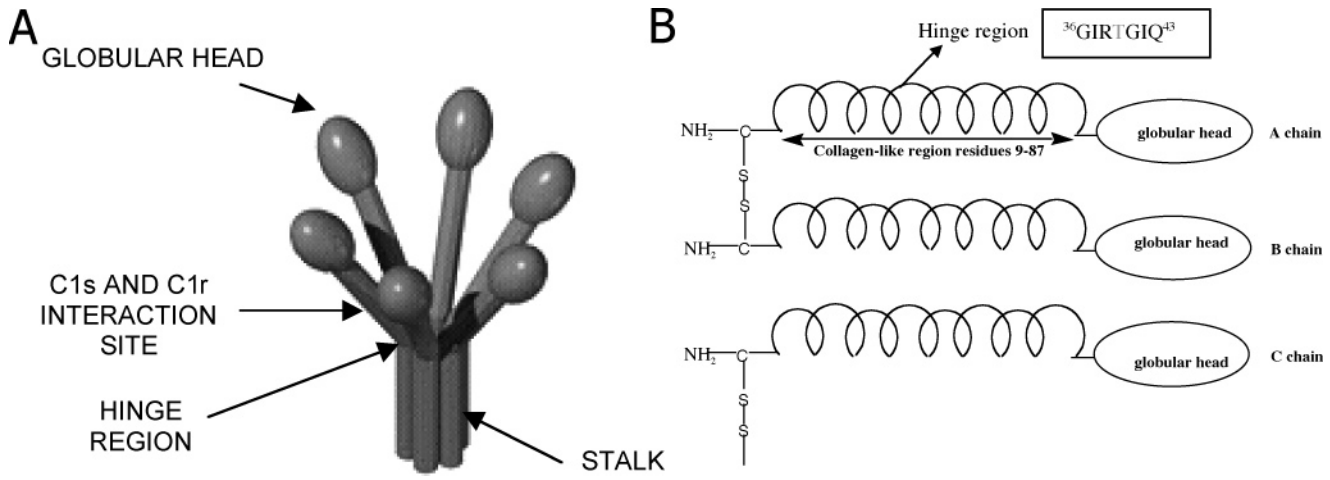

FIGURE 1: Schematic representation of the structural organization of human C1q. (A) Model showing the different structural domains and functional sites of C1q. (B) Representation of the association between the A, B, and C chains leading to formation of the six ABC heterotrimeric triple helices of C1q. The sequence responsible for the kink of the collagen-like region is shown.

tion of the tetramer occurs when C1 binds to immune complexes and is thought to involve conformational changes in C1r allowed by the flexibility of the hinge region of the CLR (*12*). The CLR has also been reported to bind a variety of nonimmune substances, leading to either activation or inhibition of C1 (*13*). Thus, binding of C1q to C-reactive protein (*14*), serum amyloid protein (*15*), or DNA (*16*) has an activating effect, whereas interaction with sulfated glycosaminoglycans (*17*−*19*), the proteoglycan dermatan sulfate decorin (*20*), or chondroitin 4-sulfate results in inhibition of the classical pathway (*21*).

We have recently shown that fucoidan binds C1q as well and its isolated collagen-like region in such a way that it prevents their interaction with the C1r moiety of C1s-C1r-C1r-C1s (*22*), yielding a functionally altered C1 complex with a decreased ability to cleave proteins C2 and C4 (*5*). The aim of this study was to obtain further insights into the fucoidan binding site(s) of the C1q molecule. On the basis of the observations that C1 assembly involves basic residues of the CLR (*23*) and that fucoidan protects lysine and arginine residues of C1q from their chemical modification (*5*), C1q was submitted to trypsin proteolysis in the presence or absence of fucoidan. In addition to providing the first analysis of C1q by mass spectrometry, this study yields evidence that the B chain moiety of the globular head is involved in the interaction with fucoidan and underscores the particular role of Arg^{109} in the charge pattern recognition property of C1q.

EXPERIMENTAL PROCEDURES

Materials. 2,5-Dihydroxybenzoic acid (DHB), α-cyano-4-hydroxycinnamic acid (HCCA), and sinapinic acid (SA), were purchased from Sigma-Aldrich (Saint-Quentin Fallavier, France). Complement protein C1q was purchased from VWR (Fontenay-sous-Bois, France). Fucoidan from the brown alga *A. nodosum* [8000 g·mol^{-1}, sulfate content 30% (w/w)] was purified as described previously (*24*), and its molecular weight was determined by high-performance size exclusion chromatography using heparin standards (*25*). Double-stranded DNA from salmon testes (type III), acrylamide-bisacrylamide (30% solution), Coomassie Brilliant Blue R, and β-mercaptoethanol were purchased from Sigma (Saint-Quentin Fallavier, France). Trypsin−EDTA and sequencing grade trypsin from bovine pancreas were purchased from Roche (Mannheim, Germany). Other chemicals and reagents were obtained from commercial sources at the highest level of purity available. All buffers were prepared using ultrapure water (milliQ, Millipore) and degassed by filtration before use.

MALDI-TOF Mass Spectrometry of C1q. Analyses were carried out on a Voyager DE-STR mass spectrometer (Applied Biosystems, Boston, MA) using two matrix solutions: (i) 60 μg/μL DHB (2,5-dihydroxybenzoic acid) in methanol and (ii) 10 μg/μL SA (sinapinic acid) in 30/70 (v/v) acetonitrile/1% formic acid. Native C1q (1 mg/mL, 15 mM sodium phosphate, pH 7.2, containing 135 mM NaCl and 40% glycerol) was diluted 5-fold in 10 mM ammonium acetate, pH 5.5, prior to being mixed with the matrix solutions, whereas trypsin-treated C1q was directly mixed with the matrix solutions. When mentioned, native or trypsin-treated C1q was reduced by dilution, 5- and 2.5-fold, respectively, in 10 mM ammonium acetate, pH 5.5, containing 1% β-mercaptoethanol, and boiled for 5 min. The solutions were loaded on the plate using the dried droplet method (*26*). Typically, mass spectra were obtained in the positive linear mode, under an acceleration voltage of 25 kV, a grid voltage of 90%, and a delay time of 1000 ns. The laser power [indicated as arbitrary units (au)] was 4000 au with the DHB matrix and 4500 au with the SA matrix.

Membrane-Assisted MALDI Analysis of Native C1q. A polyester membrane was fixed on a membrane plate (Applied Biosystems, Paris, France) using a conductive double-face adhesive (Radiospares, Beauvais, France). Undiluted native C1q (5 μL) was loaded directly on the polyester membrane. The whole system was incubated for 4 h at room temperature under moisture to avoid evaporation. After reperfusion of the residual solution, 1 μL of the SA matrix was loaded, and the resultant solution was allowed to crystallize. Under these conditions, the laser power was 4200 au.

Trypsin Digestion of C1q. C1q (5 μg) and trypsin (0.5 μg) were mixed with or without 5 μg of polyanion fucoidan or DNA in 15 μL of the digestion buffer (200 mM NH_4HCO_3, 0.5 mM $CaCl_2$, pH 8.0) and incubated for 4 h at 37 °C. The 1/10 (w/w) enzyme/protein ratio was required

to achieve proteolytic cleavage of C1q. Higher amounts of the digest were prepared for MALDI-TOF-MS analysis by incubating 20 μg of C1q with 2 μg of trypsin in 30 μL of the digestion buffer overnight at 37 °C.

SDS−PAGE Analysis. C1q was analyzed using a 13% polyacrylamide gel after heat denaturation in β-mercaptoethanol (*27*) to detect the three chains, A, B, and C. Protein bands were detected using Coomassie Blue staining. For analysis of the disulfide-bonded dimers A-B and C-C, C1q was treated under nonreducing conditions, and electrophoresis was carried out in a 7.5% acrylamide gel according to the method of Reid and Porter (*28*).

In-Gel Digestion of C1q. In-gel digestion of the three chains, A, B, and C, was performed by excision of the corresponding three discrete bands from the Coomassie Blue stained gel. After destaining in 5% acetic acid and 30% ethanol, the gel pieces were rinsed twice alternatively with 400 μL of 20 mM NH_4HCO_3, pH 8.0, and with 400 μL of 50% acetonitrile under gentle agitation (20 min each) followed by a final wash with water. The gel pieces were then freeze-dried and either stored at −20 °C or rehydrated in 30 μL of 20 mM NH_4HCO_3, pH 8.0, containing 10 μg of sequencing grade trypsin. After digestion of the samples for 5 h at 37 °C, the supernatant was collected, and the gel pieces were washed with 20 μL of 1% formic acid for 10 min at room temperature under gentle agitation. The formic acid supernatant was collected, and a last washing step with 30−40 μL of acetonitrile was performed for 10 min. All supernatants were pooled and evaporated and were then dissolved in 10 μL of 1% formic acid and desalted on C18 Zip-Tip (Millipore) according to the manufacturer's instructions. Elution of the peptides was performed with 4 μL of acetonitrile/1% formic acid (50/50 v/v), followed by 4 μL of acetonitrile/1% formic acid (80/20 v/v).

MALDI-TOF Mass Spectrometry of the C1q Tryptic Digest. The tryptic digest (0.5−1 μL) was mixed with 1 equiv of the HCCA or SA matrix. Calibration was based on peptides resulting from trypsin autolytic cleavage. Typically, mass spectra were obtained in the positive reflector mode under an acceleration voltage of 20 kV, a grid voltage of 75%, and a delay time of 225 ns. The laser power was 2900 au with the HCCA matrix and 3100 au with the SA matrix. Identification of peptides was achieved using the "MS-fit" module of the Protein Prospector program (http://prospector.ucsf.edu/ucsfhtml4.0/msfit.htm), by matching monoisotopic peptide masses in the SWISSPROT database, and using the "peptide mapping" module of the Profound program (http://129.85.19.192/profound_bin/WebProFound.exe), by matching monoisotopic peptide masses in the NCBI database. Mass tolerance was fixed at 40 ppm in both protein searches. The protein molecular weight was considered to be ±25% of the value determined by SDS−PAGE analysis. Alkylation of cysteine by acrylamide was allowed as well as one missed cleavage.

NanoESI-Q-TOF Mass Spectrometry of the C1q Tryptic Digest. Tandem mass spectrometry was carried out on a Q-STAR hybrid mass spectrometer (Applied Biosystems). The tryptic digest, diluted with an equal volume of methanol, was introduced into a glass capillary (Protana, Odense, Denmark) for nanoelectrospray ionization. Each multicharged ion, recalculated from MALDI-MS data, was submitted to fragmentation in the collision cell of the Q-TOF mass spectrometer using a collision energy between 15 and 40 eV. MS/MS sequence information was used to reconstruct peptide sequences manually. All peptide sequences were checked using BLAST (http://www.ncbi.nlm.nih.gov/blast) to confirm protein identification.

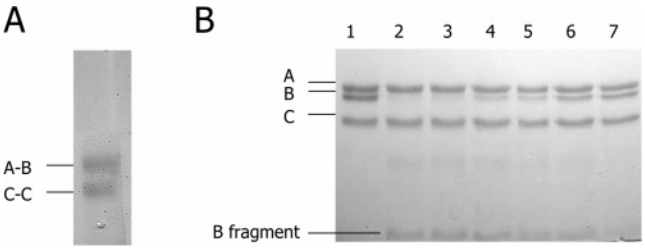

FIGURE 2: SDS−PAGE analysis of C1q. (A) Analysis of C1q under nonreducing conditions. (B) Analysis under reducing conditions. Lanes: 1, C1q (5 μg); 2, C1q (5 μg) digested with trypsin (0.5 μg); 3−7, C1q (5 μg) digested with trypsin (0.5 μg) in the presence of 0.1, 0.5, 1, 2, and 5 μg of fucoidan, respectively.

RESULTS

Electrophoretic Analysis of C1q. SDS−PAGE analysis of nonreduced C1q revealed two protein bands of apparent molecular weights 55000 and 50000 corresponding to the disulfide-bonded A-B and C-C dimers (Figure 2A), whereas reduction yielded three bands of molecular weights 31000, 29700, and 26100, corresponding to the A, B, and C chains (Figure 2B). Enzymatic deglycosylation of C1q using a combination of peptide−*N*-glycosidase F and *O*-glycosidase decreased the molecular weight of the A chain to a value close to that of the B chain (data not shown), in agreement with the location of an N-linked carbohydrate in the A chain (*29*). To confirm identification of the A, B, and C chains, the corresponding bands were digested in gel with trypsin, and the resulting fragments were analyzed by MALDI mass spectrometry, as well as by nanospray-Q-TOF sequencing for some of the peptides. From peptide mapping, a sequence coverage of approximately 25% was obtained for each chain, with high scores. Fragmentation of some peptides from each band by nanospray-Q-TOF allowed partial reconstitution of their primary sequences and unambiguous identification of A as the upper band, B as the middle band, and C as the lower band (Table 1).

Fucoidan Prevents Trypsin Cleavage of Native C1q. Native C1q was incubated with trypsin at a 1/10 (w/w) enzyme/protein ratio for 4 h at 37 °C. Unexpectedly, SDS−PAGE analysis of the reaction mixture under reducing conditions revealed that a single band, corresponding to the B chain, was modified (Figure 2B). Preliminary experiments showed that its intensity decreased when the amount of trypsin was increased. SDS−PAGE analysis also indicated that B chain consumption was concomitant with the appearance of a new band of molecular weight approximately 12000 (Figure 2B). Neither the A chain nor the C chain was cleaved, indicating selective action of trypsin on the B chain. In agreement with these results, SDS−PAGE analysis under nonreducing conditions showed that only the band corresponding to the A-B dimer decreased, the C-C dimer being unaffected (data not shown). Control experiments in which C1q was treated with a reducing agent before addition of trypsin led to complete digestion of the three chains, indicating that selective cleavage of the B chain was conformation-dependent.

Table 1: Identification of the A, B, and C Polypeptide Chains from Their 1D In-Gel Digestion by Trypsin and Mass Spectrometry Analysis

	identification scores from MALDI-TOF mass fingerprint							
			profound		protein prospector	sequencing of peptides by nanoESI-Q-TOF		
gel band	no. of matched peptides	protein coverage (%)	probability[c]	estd *Z* score[c]	MOWSE score[d]	*m/z* value of precursor ion	sequenced peptide and charge state[e]	protein identity
upper[a]	6	24[a]	1	1.88	6.20×10^4	419.9	DQPR**PAFSAIR**$^{2+}$	C1q-A
middle	7	26[b]	1	2.39	3.12×10^5	796.4	**LEQGENVFLQATDK**$^{2+}$	C1q-B
						621.9	**FDHVITNMNNNYEPR**$^{3+}$	
						383.2	**IAFSATR**$^{2+}$	
lower	5	25[b]	1	2.20	3.28×10^4	643.3	**FNAVLTN**PQGDYDTST^{3+}	C1q-C
						629.3	TNQ**VNSGGVLLR**$^{2+}$	
						542.8	FQ**SVFTVTR**$^{2+}$	
12000 band	4	20[b]	0.98	0.64	8.72×10^3	796.4	LEQGE**NVFLQATDK**$^{2+}$	C1q-B
						621.9	FDHV**ITNMNNNYEPR**$^{3+}$	

[a] Identification performed at 40 ppm, with one missed cleavage allowed, using SA as matrix. [b] Identification performed at 40 ppm, with no missed cleavage allowed, using HCCA as matrix. [c] Scores obtained from the Profound program (http://129.85.19.192/profound_bin/WebProFound.exe). The estimated *Z* score ascertains the identity of the protein when estimated $Z > 1.65$ with a probability of 1. [d] Scores obtained from the MS-fit module of the Protein Prospector program (http://prospector.ucsf.edu/ucsfhtml4.0/msfit.htm). The higher the MOWSE score, the better the identification. [e] The identified sequence of the fragmented peptide is in bold type.

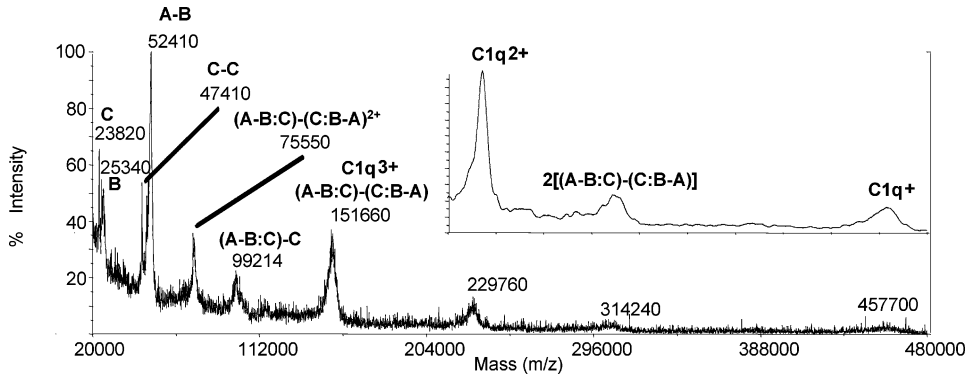

FIGURE 3: Membrane-assisted MALDI-TOF mass spectrum of native C1q. The polyester membrane was fixed on a membrane plate with a conductive double-face adhesive. Undiluted native C1q (5 μL) was loaded. After incubation for 4 h and reperfusion of the residual solution, 1 μL of the SA matrix was deposed as described under Experimental Procedures. The laser intensity was 4200 au. Insert: high mass range of the spectrum smoothed by 301 points (identical mass scale).

In a second set of experiments, tryptic cleavage of C1q was performed in the presence of two of the known C1q ligands, DNA or fucoidan (*22*). Incubation of C1q with an amount of trypsin allowing complete consumption of the B chain, in the presence of increasing amounts of fucoidan, revealed that fucoidan protected the chain B from trypsin cleavage in a dose-dependent manner (Figure 2B). Likewise, SDS−PAGE analysis under nonreducing conditions showed that the A-B dimer was protected from cleavage by fucoidan (not shown). A control experiment carried out with BSA as a protein substrate using the same 1/10 (w/w) enzyme/protein ratio showed that the proteolytic activity of trypsin was not inhibited by fucoidan. In contrast, DNA was unable to prevent tryptic cleavage of the B chain, even at high concentration (1 mg/mL).

MALDI-TOF Mass Spectrometry of C1q. Mass spectrometry analysis was employed to obtain further insights into the interaction between fucoidan and C1q. Native C1q was first analyzed by MALDI-TOF MS using two different matrices, 2,5-dihydroxybenzoic acid (DHB) and sinapinic acid (SA). SA is known to be well adapted for high molecular weight determinations (>50000) (*30*), whereas DHB gives better accuracy in the lower molecular weight range (*26*). In an initial attempt to detect the whole C1q molecule, the SA matrix was added to the protein loaded beforehand on a polyester membrane in order to withdraw glycerol from the protein solution (see Experimental Procedures) (*31*). The MALDI-TOF mass spectrum of native C1q exhibited several peaks in the high molecular weight range, with a species at *m/z* 457700 corresponding to the whole protein (Figure 3), indicating that the noncovalent interactions between the A, B, and C chains were partially retained under the experimental conditions used (*32*). Likewise, an (A-B:C)-(C:A-B) species, at *m/z* 151660 and at *m/z* 75550 for its doubly charged ion, corresponding to the so-called structural unit (*28*), was observed on the mass spectrum. Trace amounts of a species corresponding to a noncovalent dimer of this structural unit were also seen at *m/z* 314240. The peak at *m/z* 229760 was assigned to the doubly charged ion of the whole C1q complex, considering that a species comprising one and a half structural units is unlikely to form. Noncovalent association of one A-B dimer with one C-C dimer was also observed at *m/z* 99214. Finally, the mass spectrum exhibited large peaks at *m/z* 52410 and 47410, likely corresponding to A-B and C-C dimers, respectively. However, due to the altered signal/noise ratio in this range, further analysis of the low molecular weight species was achieved using a DHB matrix.

This enabled us to detect both disulfide-bonded A-B and C-C dimers at *m/z* 53312 and 48156, respectively (Figure 4), in a good agreement with the values determined by SDS−PAGE analysis under nonreducing conditions. The A, B, and C monomers were also detected at *m/z* 27635, 25683, and 24111, respectively. The peak at *m/z* 24111 assigned to the

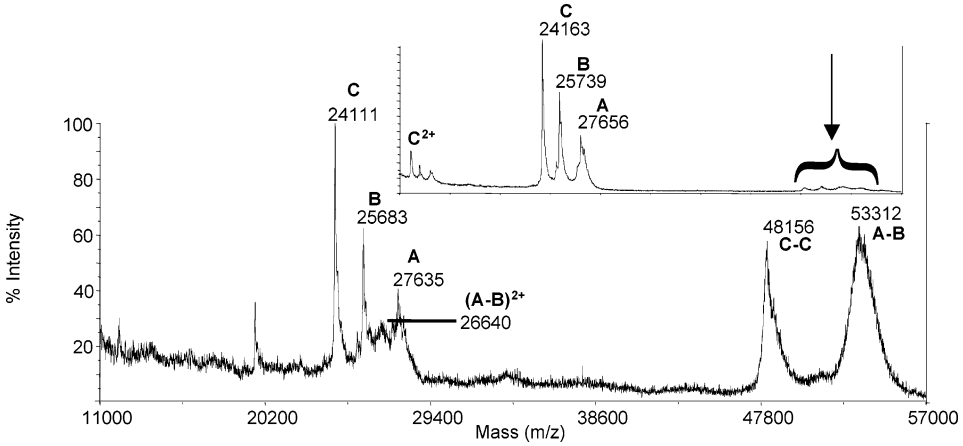

FIGURE 4: MALDI-TOF mass spectra of C1q in DHB. The spot of native C1q was prepared using the dried droplet method by mixing 1 μL of native C1q diluted 5-fold in 10 mM ammonium acetate, pH 5.5, with 1 μL of a 60 μg/μL DHB solution in methanol. The laser intensity was 4000 au. Insert: MALDI-TOF mass spectrum recorded for reduced C1q under the same conditions. The brace points out random noncovalent associations of two polypeptide chains: AA, AB, AC, BB, BC, and CC.

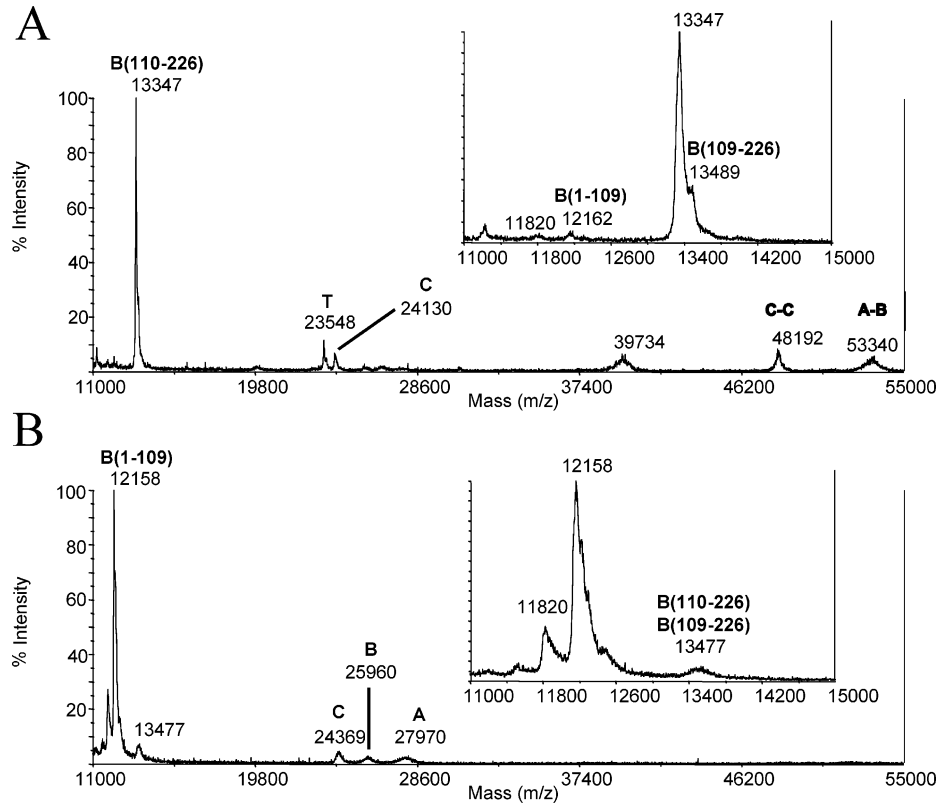

FIGURE 5: MALDI-TOF mass spectra of trypsin-treated C1q. (A) Analysis under nonreducing conditions. (B) Analysis in the presence of β-mercaptoethanol. Spectra were acquired using a DHB matrix. The laser intensity was 4000 au. T corresponds to the trypsin peak.

C chain may also correspond to the doubly charged ion of the C-C dimer, whereas the small peak at 26640 likely corresponds to the doubly charged ion of the A-B dimer. Interestingly, in addition to the expected A, B, and C chains, analysis of reduced C1q revealed a series of very weak peaks around *m*/*z* 50000 (Figure 4, insert). These peaks appear to have a distribution that could be assigned to all possible combinations between A, B, and C, suggesting trace amount formation of random associations.

MS Analysis of the C1q Tryptic Digest. The MALDI-TOF mass spectrum of trypsin-treated C1q was strikingly different from that of the native protein (Figure 5A). Thus, a major peak at *m*/*z* 13347 and an additional peak at *m*/*z* 39734, not present in untreated C1q, were observed. The summation of their *m*/*z* values yielded a mass value of 53081, consistent with that of the A-B dimer (53340). Taken together, these observations were fully consistent with the hypothesis that the two new species at *m*/*z* 13347 and 39734 arose from cleavage of the A-B dimer.

On the basis of our SDS−PAGE analysis it was assumed that the new peak at *m*/*z* 13347 was a fragment from the B chain. As this mass is about half that of the B chain, we focused on the amino acid sequence in the middle part of the B chain. This area contains several potential cleavage sites for trypsin, with in particular five neighboring arginine residues, Arg^{101}, Arg^{108}, Arg^{109}, Arg^{114}, and Arg^{129}. Among these residues, only Arg^{108} and Arg^{109} can yield C-terminal cleavage fragments with calculated masses (13412 and

13256, respectively) consistent with the peak at *m*/*z* 13347. Indeed, both of these fragments were probably formed since detailed analysis of the mass spectrum (Figure 5A, insert) revealed two components, a major one at *m*/*z* 13347 corresponding to cleavage at Arg^{109} and a minor one at *m*/*z* 13489 corresponding to cleavage at Arg^{108}. Both cleavages are expected to yield complementary fragments from the N-terminal side of the cleavage site. Few peaks of low intensity are observed in a mass range compatible with these fragments. Among them, only the peak at *m*/*z* 12162 (Figure 5A, insert) has a mass value in accordance with that of the N-terminal fragment Gln^1-Arg^{109} (calculated value = 12098). Moreover, addition of its mass value to that of the A chain accounts for the species at *m*/*z* 39734, consistent with formation of a truncated dimer between the A chain and this N-terminal fragment from the B chain. Indeed, this fragment contains the Cys^4 residue involved in the disulfide bond to the A chain, and it can be assumed that they remain covalently linked to the A chain after trypsin cleavage.

To confirm these hypotheses, MS analysis of trypsin-treated C1q was carried out in the presence of β-mercaptoethanol (Figure 5B). Under these reducing conditions, the A-B and C-C dimers, as well as the peak at *m*/*z* 39734, were not observed. This latter observation was fully consistent with the hypothesis that trypsin cleavage of the B chain yields N-terminal fragments that remain attached to the A chain through a disulfide bond. Analysis under reducing conditions also led to a strong decrease of the peak at *m*/*z* 13347, with a concomitant appearance of a new peak at *m*/*z* 12158. This latter peak was assigned to the N-terminal fragment Gln^1-Arg^{109} released from the A chain upon the reduction treatment (calculated mass = 12098). This fragment encompasses the collagen-like region of the B chain, which is known to be resistant to trypsin cleavage (*33*). Conversely, the observed strong decrease of the B chain species at *m*/*z* 13347 likely results from an increased susceptibility of this fragment to trypsin cleavage under reducing conditions, due to its high content in Arg and Lys residues. In the same way, the trypsin peak observed at *m*/*z* 23548 under nonreducing conditions (Figure 5A) was lacking under reducing conditions, likely due to autoproteolysis.

DISCUSSION

This study describes the first mass spectrometry analysis of peptides generated by in-gel digestion of the three chains of human C1q, providing unambiguous experimental evidence that these chains migrate in the order A, B, C when separated by SDS−PAGE. The three polypeptide chains are similar in length, and therefore their differential electrophoretic behavior arises in part from their respective carbohydrate content. At least 80% of the hydroxylysine residues of the collagen-like region carry an O-linked glucosylgalactosyl disaccharide (*34*). In addition, residue Asn^{124} of the A chain carries an N-linked biantennary oligosaccharide, and therefore the A chain is the most heavily glycosylated (*29*, *35*, *36*). The MALDI-MS analysis of the peptide overlapping Asn^{124} showed a mass increment of 2350 $g{\cdot}mol^{-1}$ (data not shown), in agreement with the structure previously reported for this oligosaccharide (*29*). Indeed, as listed in Table 2, the experimental mass values determined for each chain by mass spectrometry are very close to the values calculated from their amino acid content and their reported posttranslational modifications (*31*, *37*, *38*).

Table 2: Experimental *m*/*z* and Theoretical Molecular Weight Values of the Species Constituting C1q

species	av exptl mol wt ± SD (no. of measurements)	calcd mol wt[a]	difference between calcd and exptl values (%)
C	24119 ± 196 (n = 5)	23718	1.7
B	25680 ± 257 (n = 4)	25327	1.4
A	27754 ± 188 (n = 3)	27738	0.1
C-C	47919 ± 441 (n = 3)	47434	1.0
A-B	53021 ± 529 (n = 3)	53063	0.1
$(ABC)_2$	153659 ± 1746 (n = 3)	153560	0.1
C1q	459520[b]	460680	1.0
	461883 ± 4497[c]	460680	0.3
	460977 ± 5238[d]	460680	0.1

[a] Values calculated from the amino acid sequence and known posttranslational modifications (*33*, *37*−*39*). [b] Molecular weight calculated from the doubly charged ion of C1q observed on the spectrum in Figure 3. [c] Molecular weight calculated from the average experimental mass values of the A-B and C-C dimers. [d] Molecular weight calculated from the average experimental mass values of the $(ABC)_2$ structural unit.

Our MALDI-MS analyses also provide strong experimental evidence that the polypeptide chains of C1q are organized in disulfide-linked A-B and C-C dimers and for the occurrence of a noncovalently linked structural unit (A-B:C)-(C:B-A), in full agreement with the model originally proposed by Reid and Porter (*28*). These data demonstrate the ability of the membrane-assisted MALDI-TOF technique to reveal covalent and noncovalent complexes in a high molecular weight range, as exemplified by the dimeric form of the structural unit and by the intact C1q molecule observed at *m*/*z* 457700. A molecular weight of 459520 was determined from the corresponding doubly charged ion at *m*/*z* 229760, in good agreement with the theoretical molecular weight of 460680 based on amino acid sequence and carbohydrate content (*39*). Indeed, the value determined for intact C1q is likely to be misestimated due to the broadness of the peak (Figure 3) and the lack of appropriate standard for calibration in this high molecular weight range. On the other hand, the estimate derived from the mass values of the individual A, B, and C chains is probably also incorrect because the error on each measurement is multiplied by six. We therefore chose to take into account three different molecular weight estimates, deduced from the mass values of the A-B and C-C dimers, the (A-B:C)-(C:B-A) structural unit, and the doubly charged ion of C1q (Table 2). These values range from 459520 to 461883, with an average mass of 460793, very close to the theoretical value (460680). Altogether, our MS analyses fully support the hexameric model proposed for C1q by Reid and co-workers (*39*).

We have previously shown that C1q is a target protein for fucoidan (*5*) and that formation of a C1q−fucoidan complex may account for the anticomplementary activity of this polysaccharide. Because C1q is a charge pattern recognition molecule, it is assumed that interaction with its ligands mainly involves ionic interactions. Consistent with this hypothesis, binding to fucoidan was found to involve basic residues of C1q (*5*). With a view to locate these residues within the C1q molecule, we have used trypsin cleavage in the presence or absence of fucoidan. Whereas the three C1q

chains, in the reduced state, were each fully sensitive to trypsin proteolysis, unexpectedly, only the B chain was cleaved when native C1q was submitted to trypsin treatment, indicating that this highly specific cleavage requires structural integrity of the C1q molecule and/or recognition of a conformational site within the B chain. Mass spectrometry analyses allowed us to identify a C-terminal fragment at *m*/*z* 13347 and the corresponding N-terminal fragment disulfide-linked to the A chain, thereby providing clear evidence that trypsin cleavage mainly occurs at Arg^{109}, in the globular region of the B chain. The X-ray structure of the C1q globular domain has been solved (*10*), indicating that the outer surface of module B presents several exposed basic residues, including Arg^{101}, Arg^{108}, Arg^{109}, Arg^{114}, Arg^{129}, and Arg^{163}. Chemical modification and site-directed mutagenesis experiments have highlighted the involvement of Arg^{114} and Arg^{129} in the binding of IgG (*9*, *40*), but none of these studies have considered the implication of Arg^{109}. Nevertheless, the X-ray structure of the C1q globular domain (*10*) and the hydrophilicity profile of its B module (*9*) both indicate that this latter residue is one of the most exposed to the solvent. In addition, the X-ray structure shows that Arg^{109} has its guanidino group pointing outside the core of the protein, making this residue accessible and therefore available to trypsin cleavage. Compared to Arg^{109}, Arg^{114} appears to be less exposed to the surface, which may explain the lack of cleavage at this position. This may also apply to the other Arg and Lys residues present on the surface of module B, of which the side chain is possibly not fully accessible to the trypsin active site. On the other hand, the lack of cleavage in the A and C modules of the globular domain may arise from their particular surface properties (*10*) and/or from other characteristics linked to the overall C1q architecture. Thus, the fact that, unlike module B, modules A and C are likely positioned inside the C1q molecule (*12*) may limit their accessibility to trypsin. In the same way, the biantennary oligosaccharide attached to the A module likely causes steric hindrance.

Several studies have provided evidence that the B moiety of the C1q globular domain is involved in the recognition of immunoglobulins and nonimmune ligands such as the bacterial porin OmpK36, the salivary agglutinin, and the β-amyloid peptide (*10*, *41*, *42*). To check the effect of trypsin cleavage on the binding properties of C1q, trypsin-treated C1q was tested for its hemolytic activity and for its IgG binding capacity. Trypsin cleavage abolished C1q ability to interact with human IgG, as measured by surface plasmon resonance spectroscopy, and also decreased C1q hemolytic activity (data not shown). Although it may be hypothesized that trypsin cleavage disrupts the overall C1q structure, this appears unlikely since cleavage at Arg^{109} does not lead to subsequent proteolysis. A plausible explanation is therefore that cleavage at Arg^{109} alters the fine structural organization of the neighboring area, underscoring the key role of this region of the B module in the recognition of IgG.

Using affinity electrophoresis and single molecule fluorescence microscopy, we have shown previously that fucoidan binds to C1q (*5*, *22*). The present study demonstrates that interaction with fucoidan specifically prevents C1q from proteolytic cleavage by trypsin, implying that fucoidan binds the globular head of C1q, at a site close to the Arg^{109} trypsin cleavage site. It should be emphasized that this protective effect was not observed with other sulfated polysaccharides such as dextran sulfate, dermatan sulfate, and several chondroitin sulfates (data not shown), pointing out that this property is specific of fucoidan. On the other hand, we have provided evidence that fucoidan prevents interaction between C1q and C1r and thereby inhibits assembly of the C1 complex (*5*, *22*). Given the estimated distance in C1q between a globular domain and the C1r binding site (approximately 50 Å), it is unlikely that a single molecule of the low molecular weight fucoidan studied here could bind to both sites simultaneously. The most likely hypothesis is therefore that C1q binds fucoidan at two different sites, one in the globular domain and the other in the stem, in the vicinity of the interaction site with C1s-C1r-C1r-C1s. If this hypothesis is correct, fucoidan would interfere with both the binding of C1q to IgG and the assembly of C1 and, hence, would represent a very efficient tool to inhibit complement at the first step of its activation. Such a characteristic is undoubtedly of great value in terms of therapeutic application. A glycosaminoglycan chondroitin 4-sulfate proteoglycan secreted by the human B lymphocytes has been proposed to act as a physiological inhibitor of C1q through the inhibition of C1 assembly (*43*). The question arises whether this endogen-sulfated polysaccharide also interferes with IgG binding through interaction with the C1q globular head.

It is noteworthy that DNA, another polyanion, was not found to prevent C1q from trypsin cleavage. This result is consistent with the lack of DNA binding to the globular domain as observed by single molecule fluorescence microscopy (*22*). Several other studies in the literature provide support to a preferential binding of DNA to the collagen-like domain of C1q (*16*, *44*, *45*). Since DNA binding to C1q triggers C1 activation (*46*), it may be hypothesized that this involves a mechanism different from the one proposed for ligands such as IgG that interact with the globular head.

REFERENCES

1. Tissot, B., and Daniel, R. (2003) Biological properties of sulfated fucans: the potent inhibiting activity of algal fucoidan against the human Complement system, *Glycobiology 13*, 29−30.
2. Mulloy, B., Mourao, P. A., and Gray, E. (2000) Structure/function studies of anticoagulant sulphated polysaccharides using NMR, *J. Biotechnol. 77*, 123−135.
3. Daniel, R., Berteau, O., Chevolot, L., Varenne, A., Gareil, P., and Goasdoue, N. (2001) Regioselective desulfatation of sulfated L-fucopyranoside by a new sulfoesterase from the marine mollusk *Pecten maximus*, *Eur. J. Biochem. 268*, 5617−5626.
4. Chevolot, L., Mulloy, B., Ratiskol, J., Foucault, A., and Colliec-Jouault, S. (2001) A dissacharide repeat unit is the major structure in fucoidan from two species of brown algae, *Carbohydr. Res. 330*, 529−535.
5. Tissot, B., Montdargent, B., Chevolot, L., Varenne, A., Descroix, S., Gareil, P., and Daniel, R. (2003) Interaction of fucoidan with the proteins of the complement classical pathway, *Biochim. Biophys. Acta 1651*, 5−16.
6. Cooper, N. R. (1985) The classical complement pathway: activation and regulation of the first complement component, *Adv. Immunol. 37*, 151−216.
7. Kishore, U., and Reid, K. B. (2000) C1q: structure, function, and receptors, *Immunopharmacology 49*, 159−170.
8. Kaul, M., and Loos, M. (1997) Dissection of C1q capability of interacting with IgG, *J. Biol. Chem. 272*, 33234−33244.
9. Marques, G., Anton, L. C., Barrio, E., Sanchez, A., Ruiz, S., Gavilanes, F., and Vivanco, F. (1993) Arginine residues of the globular regions of human C1q involved in the interaction with immunoglobulin G, *J. Biol. Chem. 268*, 10393−10402.
10. Gaboriaud, C., Juanhuix, J., Gruez, A., Lacroix, M., Darnault, C., Pignol, D., Verger, D., Fontecilla-Camps, J. C., and Arlaud, G. J.

(2003) The crystal structure of the globular head of complement protein C1q provides a basis for its versatile recognition properties, *J. Biol. Chem. 278*, 46974−46982.
11. Roos, A., Xu, W., Castellano, G., Nauta, A. J., Garred, P., Daha, M. R., and van Kooten, C. (2004) Mini-review: A pivotal role for innate immunity in the clearance of apoptotic cells, *Eur. J. Immunol. 34*, 921−929.
12. Gaboriaud, C., Thielens, N. M., Gregory, L. A., Rossi, V., Fontecilla-Camps, J. C., and Arlaud, G. J. (2004) Structure and activation of the C1 complex of complement: unraveling the puzzle, *Trends Immunol. 25*, 368−373.
13. Gewurz, H., Ying, S. C., Jiang, H., and Lint, T. F. (1993) Non immune activators of the classical complement pathway, *Behring Inst. Mitt. 93*, 138−147.
14. Jiang, H., Robey, F. A., and Gewurz, H. (1992) Localization of sites through which C-reactive protein binds and activates complement to residues 14−26 and 76−92 of the human C1qA chain, *J. Exp. Med. 175*, 1373−1379.
15. Ying, S.-Y., Gewurz, A. T., Jiang, H., and Gewurz, H. (1993) Human serum amyloid P component oligomers bind and activate the classical complement pathway via residues 14−26 and 76−92 of the A chain collagen-like region of C1q, *J. Immunol. 150*, 169−176.
16. Schravendijk, M. R., and Dwek, R. A. (1982) Interaction of C1q with DNA, *Mol. Immunol. 19*, 1179−1187.
17. Calabrese, G. C., Recondo, M. M., Fernandez de Recondo, M. E., and Recondo, E. F. (1997) The first component of the human complement system recognizes the active fraction of heparin, *Cell Mol. Biol. 43*, 237−242.
18. Sahu, A., and Pangburn, M. K. (1993) Identification of multiple sites of interaction between heparin and the complement system, *Mol. Immunol. 30*, 679−684.
19. Weiler, J., and Linhardt, R. (1989) Comparison of the activity of polyanions and polycations on the classical and the alternative pathways of complement, *Immunopharmacology 17*, 65−72.
20. Krumdieck, R., Höök, M., Rosenberg, L., and Volanakis, J. (1992) The proteoglycan decorin binds C1q and inhibits the activity of the C1 complex, *J. Immunol. 149*, 3695−3701.
21. Silvestri, L., Baker, J., Roden, L., and Stroud, R. (1981) The C1q inhibitor in serum is a chondroitin 4-sulfate proteoglycan, *J. Biol. Chem. 256*, 7383−7387.
22. Tissot, B., Daniel, R., and Place, C. (2003) Interaction of the C1 complex of complement with sulfated polysaccharide and DNA probed by single molecule fluorescence microscopy, *Eur. J. Biochem. 270*, 4714−4720.
23. Illy, C., Thielens, N., and Arlaud, G. (1993) Chemical characterization and location of ionic interactions involved in the assembly of the C1 complex of human complement, *J. Protein Chem. 12*, 771−781.
24. Colliec, S., Boisson-Vidal, C., and Jozefonvicz, J. (1994) A low molecular weight fucoidan fraction from the brown seaweed *Pelvetia canaliculata*, *Phytochemistry 35*, 697−700.
25. Mulloy, B., Gee, C., Wheeler, S. F., Wait, R., Gray, E., and Barrowcliffe, T. W. (1997) Molecular weight measurements of low molecular weight heparins by gel permeation chromatography, *Thromb. Haemostasis 77*, 668−674.
26. Kussmann, M., Nordhoff, E., Rahbek-Nielsen, H., Haebel, S., Rossel-Larsen, M., Jakobsen, L., Gobom, J., Mirgorodskaya, E., Kroll-Kristensen, A., Palm, L., and Roepstorff, P. (1997) Matrix-assisted laser desorption/ionization mass spectrometry sample preparation techniques designed for various peptide and protein analytes, *J. Mass Spectrom. 32*, 593−601.
27. Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4, *Nature 227*, 680−685.
28. Reid, K., and Porter, R. R. (1976) Subunit composition and structure of subcomponent C1q of the first component of human complement, *Biochem. J. 155*, 19−23.
29. Mizuochi, T., Yonemasu, K., Yamashita, K., and Kobata, A. (1978) The asparagine-linked sugar chains of subcomponent C1q of the first component of human complement, *J. Biol. Chem. 253*, 7404−7409.
30. Moniatte, M., Lesieur, C., Vecsey-Semjen, B., Buckley, J. T., Pattus, F., Van der Goot, F. G., and Van Dorsselaer, A. (1997) Matrix-assisted laser-desorption−ionization time-of-flight mass spectrometry in the subunit stoichiometry study of high mass non-covalent complexes, *Int. J. Mass Spectrom. Ion Processes 169−170*, 179−199.
31. Blackledge, J. A., and Alexander, A. J. (1995) Polyethylene membrane as a sample support for direct matrix-assisted laser desorption/ionization mass spectrometry analysis of high mass proteins, *Anal. Chem. 67*, 843−848.
32. Rosinke, B., Strupat, K., Hillenkamp, F., Rosenbusch, J., Dencher, N., Krüger, U., and Galla, H.-J. (1995) Matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS) of membrane proteins and non-covalent complexes, *J. Mass Spectrom. 30*, 1462−1468.
33. Reid, K. (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.
34. Shinkai, I., and Yonemasu, K. (1979) Hydroxylysine-linked glycosides of human complement subcomponent C1q and various collagens, *Biochem. J. 177*, 847−852.
35. Reid, K. B., Lowe, D. M., and Porter, R. R. (1972) Isolation and characterization of C1q, a subcomponent of the first component of complement, from human and rabbit sera, *Biochem. J. 130*, 749−763.
36. Reid, K. B. (1974) A collagen-like amino acid sequence in a polypeptide chain of human C1q (a subcomponent of the first component of complement), *Biochem. J. 143*, 189−203.
37. Reid, K. B., and Thompson, E. O. (1978) Amino acid sequence of the N-terminal 108 amino acid residues of the B chain of subcomponent C1q of the first component of human complement, *Biochem. J. 173*, 863−868.
38. Reid, K., Gagnon, J., and Frampton, J. (1982) Completion of the amino acid sequences of the A and B chains of the subcomponent C1q of the first component of human complement, *Biochem. J. 203*, 559−569.
39. Reid, K. B. (1983) Proteins involved in the activation and control of the two pathways of human complement, *Biochem. Soc. Trans. 11*, 1−12.
40. 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., and 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.
41. Kojouharova, M. S., Tsacheva, I. G., Tchorbadjieva, M. I., Reid, K. B., and Kishore, U. (2003) Localization of ligand-binding sites on human C1q globular head region using recombinant globular head fragments and single-chain antibodies, *Biochim. Biophys. Acta 1652*, 64−74.
42. Tacnet-Delorme, P., Chevallier, S., and Arlaud, G. J. (2001) β-amyloid fibrils activate the C1 complex of complement under physiological conditions: evidence for a site for Aβ on the C1q globular regions, *J. Immunol. 167*, 6374−6381.
43. Kirschfink, M., Blase, L., Engelmann, S., and Schwartz-Albeiz, R. (1997) Secreted chondroitin sulfate proteoglycan of the human B cell lines binds to the complement protein C1q and ihnibits complex formation of C1, *J. Immunol. 158*, 1324−1331.
44. Rosenberg, A. M., Prokopchuk, P. A., and Lee, J. S. (1988) The binding of native DNA to the collagen-like segment of C1q, *J. Rheumatol. 15*, 1091−1096.
45. Uwatoko, S., and Mannik, M. (1990) The location of binding sites on C1q for DNA, *J. Immunol. 144*, 3484−3488.
46. Jiang, H., Cooper, B., Robey, F. A., and Gewurz, H. (1992) DNA binds and activates complement via residues 14−26 of the C1q 1 chain, *J. Biol. Chem. 267*, 25597−25601.

BI047802H