Structure of the Catalytic Region of Human Complement Protease C1̄s: Study by Chemical Cross-Linking and Three-Dimensional Homology Modeling†

Véronique Rossi,‡ Christine Gaboriaud,§ Monique Lacroix,‡ Jacques Ulrich,|| Juan Carlos Fontecilla-Camps,§ Jean Gagnon,‡ and Gérard J. Arlaud*,‡

Laboratoire d'Enzymologie Moléculaire, Laboratoire de Cristallogénèse et Cristallographie des Protéines, and Laboratoire de Spectrométrie de Masse des Protéines, Institut de Biologie Structurale Jean-Pierre Ebel (CEA-CNRS), 41, avenue des Martyrs, 38027 Grenoble Cedex 1, France

Received December 1, 1994; Revised Manuscript Received January 26, 1995⊗

ABSTRACT: C1̄s is a multidomain serine protease that is responsible for the enzymic activity of C1̄, the first component of the classical pathway of complement. Its catalytic region (γ-B) comprises two contiguous complement control protein (CCP) modules, IV and V (about 60 residues each), a 15-residue intermediary segment, and the B chain (251 residues), which is the serine protease domain. With a view to identify domain−domain interactions within this region, the γ-B fragment of C1̄s, obtained by limited proteolysis with plasmin, was chemically cross-linked with the water-soluble carbodiimide 1-ethyl-3-[3-(dimethylamino)propyl]carbodiimide; then cross-linked peptides were isolated after CNBr cleavage and thermolytic digestion. N-Terminal sequence and mass spectrometry analyses allowed us to identify two cross-links between Lys 405 of module V and Glu 672 of the B chain and between Glu 418 of the intermediary segment and Lys 608 of the B chain. Three-dimensional modeling of the CCP modules IV and V and of the catalytic B chain was also carried out on the basis of their respective homology with the 16th and 5th CCP modules of complement factor H and type I serine proteases. The information provided by both the chemical cross-linking studies and the homology modeling enabled us to construct a three-dimensional model for the assembly of the C-terminal part of the γ-B region, comprising module V, the intermediary segment, and the B chain. This model shows that module V interacts with the serine protease B chain on the side opposite to both the activation site and the catalytic site. Functional implications of this interaction are discussed in terms of the possible role of module V in the specific recognition and positioning of C4, one of the two substrates of C1̄s.

The catalytic subunit of human C1, the first component of the classical pathway of complement, is a Ca^{2+}-dependent tetrameric association (C1s−C1r−C1r−C1s) of two homologous multidomain serine proteases, C1r and C1s, which undergo sequential activation when C1 binds, through its third subcomponent C1q, to immune complexes or various nonimmune activators. Autoactivation of C1r yields an active enzyme, C1̄r, which in turn converts proenzyme C1s into C1̄s. The latter is a highly specific protease with trypsin-like specificity that is responsible for limited proteolysis of C4 and C2, the next components of the complement cascade [reviewed by Cooper (1985), Schumaker et al. (1987), and Arlaud et al. (1987)].

Human proenzyme C1s, a 673-residue polypeptide (Tosi et al., 1987; Mackinnon et al., 1987), is cleaved upon activation between Arg 422 and Ile 423 to yield two disulfide-linked chains, A (N-terminal) and B (C-terminal) (Spycher et al., 1986). C1̄s is thought to be organized in two functional regions: (i) an interaction region (α) corresponding to the N-terminal half of the A chain, responsible for Ca^{2+}-dependent C1r−C1s interactions involved in assembly of the C1s−C1r−C1r−C1s tetramer; (ii) a catalytic region (γ-B) responsible for the enzymic activity of C1̄, comprising the C-terminal part of the A chain (γ) and the B chain, connected to each other through a single disulfide bridge. The B chain (251 residues), which contains the active site, is homologous to the catalytic domains of type I serine proteases. The γ region comprises two "complement control protein" (CCP)[1] modules, IV and V (about 60 residues each), homologous to those mostly found, usually in multiple copies, in complement regulatory proteins (Reid et al., 1986). Module V bears, at position 391, a complex-type N-linked oligosaccharide (Y. Petillot, P. Thibault, N. M. Thielens, V. Rossi, M. Lacroix, B. Coddeville, G. Spik, V. N. Schumaker, J. Gagnon, and G. J. Arlaud, unpublished experiments). It is followed by a 15-residue intermediary segment which in proenzyme C1s connects module V to the B chain moiety and is C-terminal to the γ region in C1̄s.

The γ-B fragment, which is obtained by limited proteolysis of C1̄s with plasmin, is visualized by electron microscopy after rotary shadowing as a compact structure (Villiers et

† This is Publication No. 254 of the Institut de Biologie Structurale Jean-Pierre Ebel. This work was supported by the Commissariat à l'Energie Atomique and the Centre National de la Recherche Scientifique.

* To whom correspondence should be addressed.
‡ Laboratoire d'Enzymologie Moléculaire.
§ Laboratoire de Cristallogénèse et Cristallographie des Protéines.
|| Laboratoire de Spectrométrie de Masse des Protéines.

⊗ Abstract published in *Advance ACS Abstracts,* May 1, 1995.

[1] Abbreviations: CCP, complement control protein module; EDC, 1-ethyl-3-[3-(dimethylamino)propyl]carbodiimide; EDTA, ethylenediaminetetraacetic acid; FAB, fast atom bombardment; HPLC, high-pressure liquid chromatography; PTH, phenylthiohydantoin; SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis. The nomenclature of complement components is that recommended by the World Health Organization; activated components are indicated by an overhead bar, e.g., C1̄s.

0006-2960/95/0434-7311$09.00/0

al., 1985). However, neutron scattering studies indicate that this fragment has a relatively large radius of gyration, not compatible with a single globular domain (Zaccaï et al., 1990). Moreover, studies by differential scanning calorimetry (Medved et al., 1989) give evidence for the presence of three independently folded domains, corresponding to the CCP modules IV and V and the B chain. From a functional point of view, studies indicate that the γ region may be involved in the binding of C4 (Matsumoto & Nagaki, 1986; Matsumoto et al., 1989). However, apart from homology modeling of the catalytic B chain (Carter et al., 1984), no information is available yet on the three-dimensional structure of the individual domains of the $C\bar{1}s$ γ-B region and on the assembly of this region.

With the view to identify domain−domain interactions within the catalytic region of $C\bar{1}s$, the γ-B fragment was submitted to chemical cross-linking with EDC, a zero-length cross-linking agent able to convert ionic bonds between carboxyl groups and ϵ-amino groups into covalent pseudopeptide bonds (Means & Feeney, 1971). At the same time, three-dimensional modeling of the individual domains from γ-B was also performed on the basis of their homology with proteins of known structures, namely, type I serine proteases and CCP modules. In a second step, the complementary information provided by both the chemical cross-linking and the homology modeling was utilized to construct a three-dimensional model of the C-terminal part of the γ-B region, including the CCP module V, the 15-residue intermediary segment, and the B chain.

EXPERIMENTAL PROCEDURES

Materials. Human plasmin was obtained from Kabi Vitrum, Stockholm, Sweden. Thermolysin from *Bacillus thermoproteolyticus* was from Boehringer Mannheim France S.A. EDC and diisopropyl fluorophosphate were purchased from Sigma.

Isolation of Fragment $C\bar{1}s$ γ-B. $C\bar{1}s$ was purified from human plasma as described previously (Arlaud et al., 1979) and subjected to limited proteolysis with plasmin as described by Thielens et al. (1990). Purification of the $C\bar{1}s$ γ-B fragment was realized by high-pressure hydrophobic interaction chromatography on a TSK-Phenyl 5PW column (7.5 mm × 75 mm) (LKB) as described previously (Thielens et al., 1990), except that the plasmin digest was dialyzed against 1.5 M $(NH_4)_2SO_4$/0.1 M Na_2HPO_4 (pH 7.4) prior to loading onto the column. Alternatively, purification was achieved by ion-exchange chromatography on a Mono Q HR 5/5 column (Pharmacia). In that case, the plasmin digest was dialyzed against 20 mM Na_2HPO_4 (pH 7.4) and loaded onto the column equilibrated in the same buffer, and elution was carried out by a linear NaCl gradient from 0 to 0.25 M in 20 min. The flow rate was 0.5 mL/min, and fragment γ-B was detected from its absorbance at 280 nm. The concentrations of purified $C\bar{1}s$ and its γ-B fragment were determined from A_{280} measurements by using values of E (1%, 1 cm) = 14.5 and 18.3, respectively (Thielens et al., 1990; Zaccai et al., 1990). The average molecular weights of $C\bar{1}s$ (79 800) and $C\bar{1}s$ γ-B (47 520) were determined experimentally by electrospray ionization mass spectrometry (Y. Petillot, P. Thibault, N. M. Thielens, V. Rossi, M. Lacroix, B. Coddeville, G. Spik, V. N. Schumaker, J. Gagnon, and G. J. Arlaud, unpublished experiments).

SDS−PAGE Analysis. Proteins were analyzed on 12.5% polyacrylamide gels as described by Laemmli (1970). Phosphorylase *b* (M_r 94 000), bovine serum albumin (M_r 67 000), ovalbumin (M_r 43 000), carbonic anhydrase (M_r 30 000), soybean trypsin inhibitor (M_r 20 100), and α-lactalbumin (M_r 14 400), all reduced and alkylated, were used as molecular weight markers. Protein staining was performed with Coomassie Blue, and gels were scanned using a Shimadzu Model CS 9000 gel scanner.

Chemical Cross-Linking. The isolated γ-B fragment (160 nmol) was dialyzed against 150 mM NaCl/20 mM 2-(*N*-morpholino)ethanesulfonic acid (pH 6.0), diluted to 0.2 mg/mL, and then incubated for 2 h at 30 °C in the presence of 10 mM EDC. The reaction mixture was dialyzed exhaustively against the same buffer without EDC and then against 0.5% acetic acid and freeze-dried. The cross-linked products were dissolved in 8 mL of 6 M guanidine hydrochloride/2 mM EDTA/0.4 M Tris-HCl (pH 8.5). Reduction of the disulfide bridges was achieved by incubation in the presence of 0.5% (v/v) 2-mercaptoethanol for 2 h at 25 °C, and alkylation was performed by a further incubation for 2 h at 25 °C in the presence of 4% (v/v) 4-vinylpyridine. The reduced and S-pyridylethylated material was dialyzed against 0.5% acetic acid and then freeze-dried. After dissolution in 6 M urea/0.2 M formic acid (2.6 mL), the cross-linked products were fractionated by high-pressure gel permeation on a TSK G-3000 SW column (7.5 mm × 600 mm) (LKB) equilibrated in the same medium. The flow rate was 1 mL/min, and peaks were detected by absorbance at 280 nm. Fractions were dialyzed against 0.5% acetic acid and freeze-dried.

CNBr Cleavage of the Cross-Linked γ-B Fragment. The reduced and S-pyridylethylated, cross-linked γ-B fragment (48 nmol) was dissolved in 70% formic acid (1.2 mL) containing CNBr (65 mmol) and kept in the dark for 24 h at 4 °C. The same treatment was applied to a reduced and S-pyridylethylated, non-cross-linked γ-B sample (31 nmol). After dilution 1:15 with water, the CNBr-cleavage peptides from both γ-B samples were freeze-dried, redissolved in 0.1% trifluoroacetic acid, and then subjected to high-pressure gel permeation on TSK G-3000 SW and TSK G-2000 SW columns (7.5 mm × 600 mm) equilibrated in 0.1% trifluoroacetic acid and used together in order of decreasing pore size, as described by Carter et al. (1983). The flow rate was 1 mL/min, and peptides were detected by absorbance at 215 nm. Fractions 1 and 3 from control γ-B and fraction X1 from cross-linked γ-B were further fractionated by reverse-phase HPLC on a C4 Deltapak column (3.9 mm × 150 mm) (Waters Associates). The column was equilibrated with a mixture of solutions A (0.1% trifluoroacetic acid) and B (acetonitrile/methanol/propan-2-ol, 1:1:1 v/v/v) in the ratio 95:5 (v/v) and then eluted with a linear gradient to give a final ratio of 20:80 (v/v) in 60 min. The flow rate was 0.8 mL/min, and peptides were detected from absorbance at 215 nm.

Thermolytic Cleavage of the Cross-Linked Peptides. Cross-linked CNBr-cleavage peptides contained in fraction X1A (2−3 nmol) were dissolved in 0.5 mL of 2 mM $CaCl_2$/0.1 M NH_4HCO_3 (pH 8.0) and digested by 5% (w/w) thermolysin for 4 h at 60 °C. After the reaction, the pH of the mixture was adjusted to 2.0 with trifluoroacetic acid, and the thermolytic peptides were separated by reverse-phase HPLC on a C4 column as described above.

N-Terminal Sequence Analysis. N-Terminal sequence analyses were performed using an Applied Biosystems Model 477A protein sequencer, and amino acid PTH derivatives were identified and quantitated on-line with a Model 120A HPLC system, as recommended by the manufacturer. Proteins separated by SDS−PAGE were electrophoretically transferred to polyvinylidene difluoride membranes (Matsudaira, 1987), according to the protocol defined for ProBlott membranes (Applied Biosystems).

Mass Spectrometry Analyses. FAB mass spectrometry analyses were carried out on a VG analytical ZAB-SE double-focusing mass spectrometer, as described previously (Thielens et al., 1990). Peptides were dissolved in 5% acetic acid, and thioglycerol was used as the matrix.

Electrospray ionization mass spectra were obtained on an API III triple−quadrupole mass spectrometer (PE/Sciex, Thornhill, Ontario, Canada) equipped with a nebulizer-assisted electrospray (ionspray) source. The mass spectrometer was scanned from *m/z* 400 to 1800, with steps of 0.5 or 1 *m/z* unit. The dwell time was 2 ms, and the resolution was 1 mass unit. For each scan, an average of 5−10 scans were accumulated. Samples submitted to flow injection analysis were dissolved in methanol/water/acetic acid (25:74:1) (v/v/v) and introduced by means of a syringe pump into a Valco C6W injector. For liquid chromatography/mass spectrometry analyses, samples were loaded onto an home-made capillary column (Nucleosil C8, 0.25 mm × 150 mm) and eluted at a flow rate of 5 μL/min with a linear gradient of 5−90% acetonitrile in 0.1% trifluoroacetic acid over 60 min. The effluent was introduced directly into the electrospray source via a 50 μm × 70 cm fused silica capillary.

Peptide Synthesis. Peptide Q665−D673 was synthesized chemically by the stepwise solid-phase method (Barany & Merrifield, 1980) in an Applied Biosystems 430A automated synthesizer, using the *tert*-butyloxycarbonyl strategy. Deprotection of the side chains and cleavage of the peptide from the resin were performed by HF treatment (Tam et al., 1983). Purification was achieved by preparative reverse-phase HPLC on a 300 Å Deltapak C18 column (1.9 cm × 30 cm) (Waters Associates), and the mass of the purified peptide was checked by FAB mass spectrometry.

Computer-Assisted Three-Dimensional Homology Modeling. The program O (Jones et al., 1991) was used on an ESV/20 graphics station (Evans & Sutherland) to construct homology-based three-dimensional models of the protein modules of the γ-B fragment and to monitor their assembly interactively on the basis of the information provided by chemical cross-linking.

The protocol used for modeling the individual modules was similar to that described in detail by Greer (1990) in the case of mammalian serine proteases. A set of homologous three-dimensional structures is used as a reference template to follow the local degree of structural conservation or divergence (mainly found at surface loops and turns) within the family. The superposition of these reference structures is performed on a set of highly conserved residues (such as the catalytic triad in the case of serine proteases) and then refined iteratively on all equivalent Cα positions (distant by less than 3.5 Å in a first step and by less than 1.5 Å in a second step). The sequence of the protein to be modeled is aligned onto the sequences of the reference set, avoiding, as much as possible, insertions or deletions in the conserved structural core of the reference set of structures. The reference structure exhibiting the fewest insertions and deletions relative to the model to be constructed is then used as the primary template. The nonidentical residues of this template within the conserved regions are computationally mutated to the corresponding residues of the model. In areas where the three-dimensional structures of the reference set diverge, a single reference structure is used as a template if the number of residues is equivalent, selecting when possible the one with higher sequence identity. The remaining areas, where the lengths of the model loops differ from those in the reference structures, are computationally constructed using a set of loops of the same length from a database of high-resolution X-ray structures (Jones & Thirup, 1986), chosen according to the similarity to the reference set of structures. The Cα three-dimensional models obtained by this method were not energy-minimized.

The serine proteases used as reference structures for modeling the B chain were the following: bovine chymotrypsin (Birktoft & Blow, 1972); bovine trypsin (Bartunik et al., 1989); porcine elastase (Sawyer et al., 1978); porcine kallikrein (Bode et al., 1983); rat mast cell protease (Remington et al., 1988); *Streptomyces griseus* trypsin (Read & James, 1988); rat tonin (Fujinaga & James, 1987); human leucocyte elastase (Navia et al., 1989); bovine chymotrypsinogen (Wang et al., 1985). The structures chosen were all determined at a resolution equal to or better than 1.8 Å, and coordinates were obtained from the Protein Data Bank (Abola et al., 1987; Bernstein et al., 1977). The CCP modules IV and V were modeled from the coordinates of the homologous 5th, 15th, and 16th modules from human factor H (Norman et al., 1991; Barlow et al., 1992, 1993).

RESULTS

Isolation of Fragment γ-B and Chemical Cross-Linking with EDC. Human C1̄s was submitted to limited proteolysis with plasmin, and purification of the γ-B fragment from the plasmin digest was carried out by either high-pressure hydrophobic interaction chromatography or ion-exchange chromatography, as described under Experimental Procedures. In both cases, fragment γ-B was eluted as a single homogeneous peak, and the purification yield was about 25%. Analysis of the purified γ-B fragment by SDS−PAGE under nonreducing conditions showed that it migrated as a single band of M_r approximately 55 000. Reduction and alkylation yielded the B chain (M_r 31 000) and the γ fragment, which was usually resolved into four bands with M_r values ranging from 24 000 to 26 000. N-Terminal sequence analysis of the material contained in these bands after SDS−PAGE analysis and electrotransfer onto polyvinylidene difluoride showed a single major sequence, Leu-Arg-Tyr-His-Gly..., indicating that all observed γ species arose from cleavage of the Lys−Leu bond at position 269 of the C1̄s A chain, in agreement with previous findings (Thielens et al., 1990).

Preliminary cross-linking studies of the γ-B fragment with EDC were performed at pH 6.5 and at low protein concentration (0.2 mg/mL), in order to minimize intermolecular cross-links. Under these conditions, as judged from SDS−PAGE analysis under reducing conditions, incubation of γ-B for 2 h at 30 °C in the presence of increasing concentrations (1−20 mM) of EDC led to the progressive disappearance of both the B chain and the γ fragment. Concomitantly, the

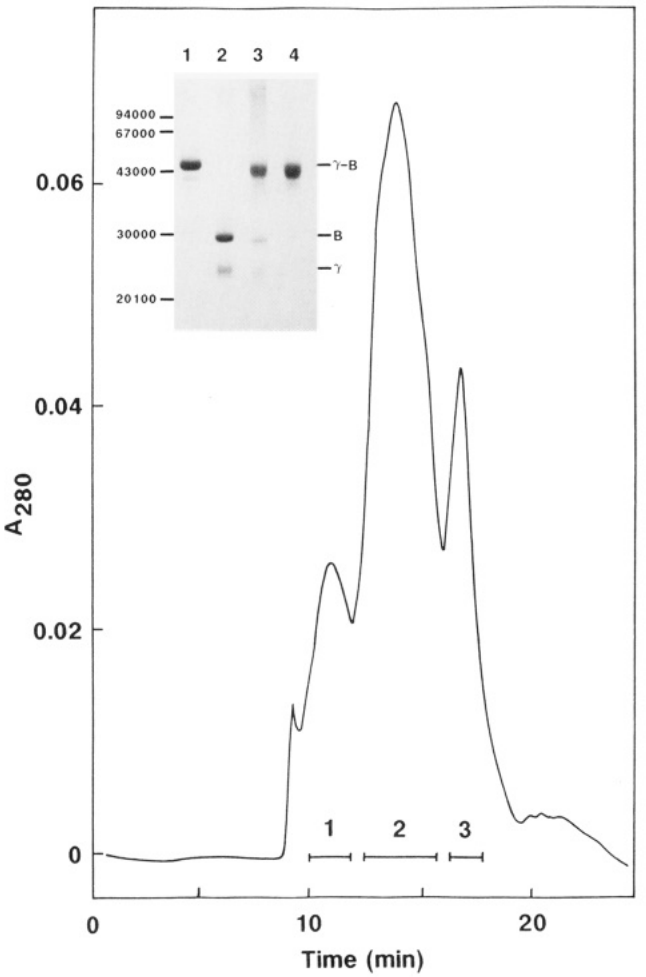

FIGURE 1: Isolation of the cross-linked γ-B species by high-pressure gel permeation. Fragment γ-B was cross-linked with EDC, reduced with 2-mercaptoethanol, and alkylated with 4-vinylpyridine as described under Experimental Procedures. The reduced and S-pyridylethylated, cross-linked material was fractionated by high-pressure gel permeation on a TSK G-3000 SW column equilibrated in 6 M urea/0.2 M formic acid. Proteins were detected by absorbance at 280 nm, and pools were made as indicated by bars. The insert shows SDS−PAGE analysis of (lane 1) the native γ-B fragment, unreduced; (lane 2) the native γ-B fragment, reduced; (lane 3) the reduced and S-pyridylethylated, cross-linked material loaded onto the column; and (lane 4) the cross-linked γ-B species isolated from fraction 2.

reaction yielded a first cross-linked γ-B species of apparent M_r 54 000, which in turn was progressively converted into a second species of apparent M_r 50 000, suggesting the occurrence of a two-step process. Several species of higher molecular weight (80 000−150 000), likely resulting from intermolecular cross-linking, also appeared at EDC concentrations above 10 mM. Decreasing the pH from 6.5 to 6.0 significantly increased the yield of the M_r 50 000 cross-linked species which, under the reaction conditions used for preparative experiments (see Experimental Procedures), reached 50−60% of the total species, as estimated from gel scanning.

One hundred sixty nanomoles of the γ-B fragment was submitted to chemical cross-linking with EDC, then reduced with 2-mercaptoethanol, and alkylated with 4-vinylpyridine, as described under Experimental Procedures. The reduced and S-pyridylethylated, cross-linked material was then fractionated by high-pressure gel permeation under denaturing conditions, as illustrated in Figure 1. This allowed us

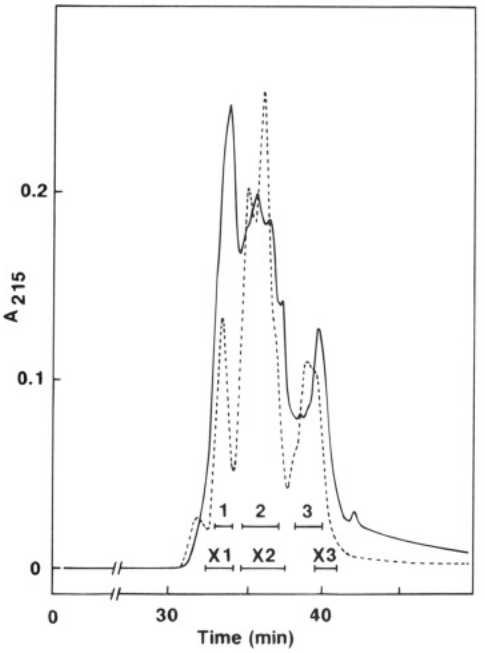

FIGURE 2: Initial fractionation by high-pressure gel permeation of the CNBr-cleavage peptides from cross-linked γ-B and from control γ-B. The CNBr-cleavage peptides from cross-linked γ-B (continuous line) and from control γ-B (dashed line) were separated on TSK G-3000 SW and TSK G-2000 SW columns as described under Experimental Procedures. Peptides were detected by absorbance at 215 nm, and pools were made as indicated by bars. Fractions 1, 2, and 3 refer to control γ-B, and fractions X1, X2, and X3 refer to cross-linked γ-B.

to separate the cross-linked γ-B species, contained in pool 2, from the above-described species of higher molecular weights, eluted in pool 1, and from the residual B chain and γ fragment, both found in pool 3. Although the cross-linked γ-B species obtained at this stage contained traces of the B chain, as shown on Figure 1, this material was used for the studies described below without further purification.

Isolation of Cross-Linked Peptides. In order to isolate cross-linked peptides containing cross-linking sites between the γ fragment and the B chain, the reduced and S-pyridylethylated, cross-linked γ-B species was submitted to CNBr cleavage as described under Experimental Procedures. The same treatment was also applied to a reduced and S-pyridylethylated but non-cross-linked control γ-B sample, and the CNBr-cleavage peptides from both γ-B samples were initially fractionated by high-pressure gel permeation, as illustrated in Figure 2. From the number of methionine residues present in the γ fragment and the B chain, CNBr cleavage was expected to yield 12 peptides (Figure 3).

As shown in Figure 2, the peptides from the control γ-B sample were separated into three major fractions, the contents of which were determined by N-terminal sequence analysis and electrospray ionization mass spectrometry. Fraction 1 only contained peptide E379−R422, the C-terminal peptide from fragment γ. Three peptides from the B chain (Y471−M487, L488−M526, D546−M601) and peptide P278−M378 from γ were found in fraction 2. The last fraction, 3, contained five peptides from the B chain (G527−M545, I602−M610, D611−M661, K622−M664, Q665−D673) and peptide L270−M277 from γ. For unknown reasons, peptide I423−M470, the N-terminal peptide from the B chain, could not be recovered from the column.

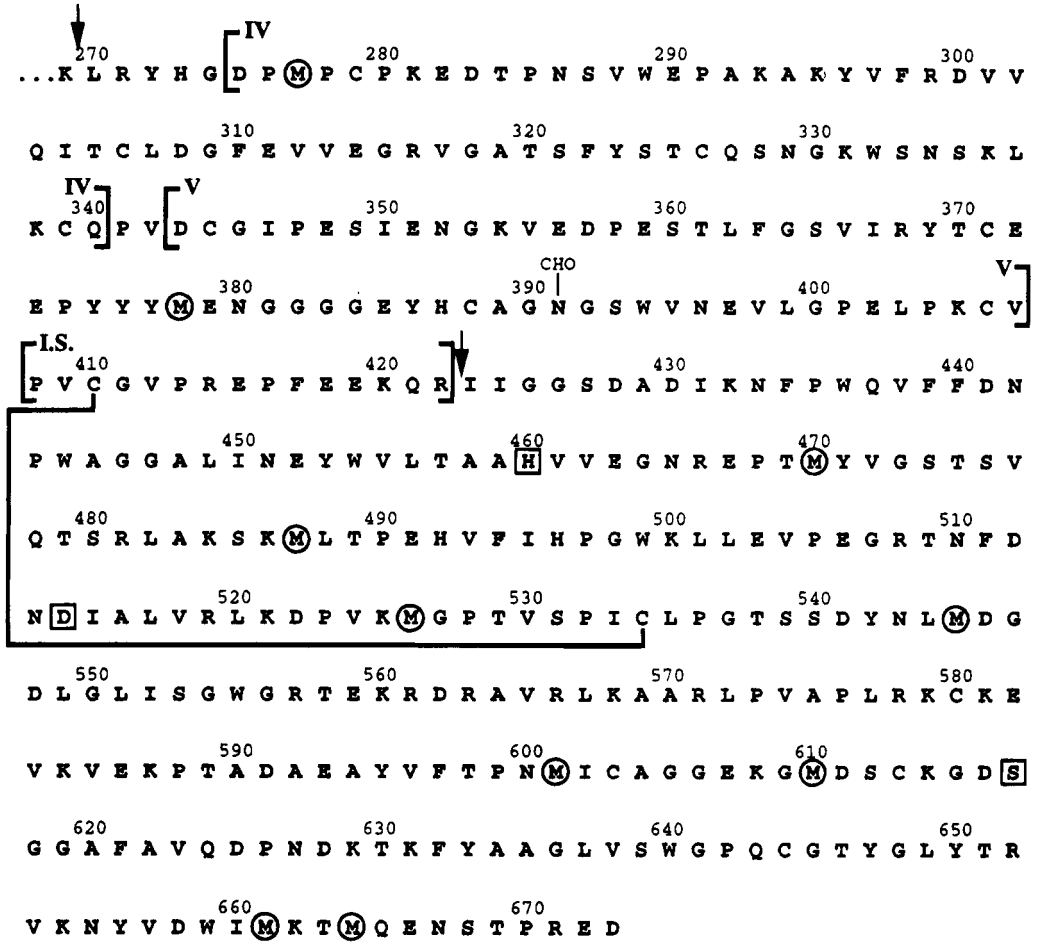

FIGURE 3: Amino acid sequence of the C1s γ-B fragment. The amino acid numbering is that of intact C1s (Mackinnon et al., 1987; Tosi et al., 1987). The site of plasmin cleavage generating the γ-B fragment and the activation site of C1s defining fragment γ and the B chain are indicated by arrows. Methionine residues are circled, and active site residues are squared. The boundaries of the CCP modules IV and V and of the intermediary segment (I.S.) are indicated as defined in this study. The only disulfide bridge shown is that connecting, through half-cystine residues 410 and 534, the γ fragment to the B chain (Hess et al., 1991). CHO indicates the N-linked oligosaccharide.

Separation of the CNBr-cleavage peptides from the cross-linked γ-B sample by the same method also yielded three major fractions, but the elution profile showed marked differences (Figure 2). In particular, the first fraction, X1, was significantly increased, whereas noticeable differences in the shape and/or elution position of the following fractions, X2 and X3, were also observed. N-Terminal sequence analysis of fraction X1 indicated that, in addition to peptide E379–R422 also found in the corresponding fraction 1, it contained equivalent amounts of peptides I602–M610 and Q665–D673, as well as lower amounts of peptides P278–M378 and L488–M526. The presence of the latter two peptides is likely a consequence of contamination by the neighboring fraction X2, which showed essentially the same peptide contents as fraction 2. In contrast, fraction X3 was found to lack peptides I602–M610 and Q665–D673, identified in the corresponding fraction 3.

The above observations strongly suggested that fraction X1 contained cross-linked material, involving at least peptides I602–M610 and Q665–D673 from the B chain. Therefore, this was further analyzed by reverse-phase HPLC, yielding two major fractions, as shown in Figure 4. N-Terminal sequence analysis of fraction X1A yielded three sequences corresponding to peptides E379–R422, I602–M610, and Q665–D673. Although this fraction clearly showed two components, separate analyses indicated that both contained equivalent proportions of the three peptides. Comparative chromatographic analysis by the same method of the corresponding peptides obtained from CNBr cleavage of the non-cross-linked γ-B sample indicated that peptide E379–R422 normally eluted later than fraction X1A (see Figure 4), whereas peptides I602–M610 and Q665–D673 both eluted in the injection peak. It appeared likely, therefore, that fraction X1A contained cross-linked material involving these three peptides.

N-Terminal sequence analysis of the heterogeneous fraction X1B yielded two major sequences, corresponding to peptides P278–M378 and L488–M526. As shown in Figure 4, the corresponding peptides obtained from the untreated γ-B sample eluted on both sides of peak X1B. Analysis of fraction X1B by electrospray ionization mass spectrometry showed no significant *m/z* peak attributable to peptide L488–M526 but revealed a coherent series of peaks with mass values differing from the expected mass of peptide P278–M378 by increments of 28 and/or 156, corresponding respectively to the addition of a formyl group and to the fixation of an EDC molecule. It appeared likely, therefore, that fraction X1B contained various formylated and EDC-modified forms of peptide P278–M378, coexisting with, rather than cross-linked to, peptide L488–M526, the latter being probably also chemically modified. Although the presence of minor cross-linked species in fraction X1B could not be entirely ruled out, further investigations were conducted solely on the material contained in fraction X1A.

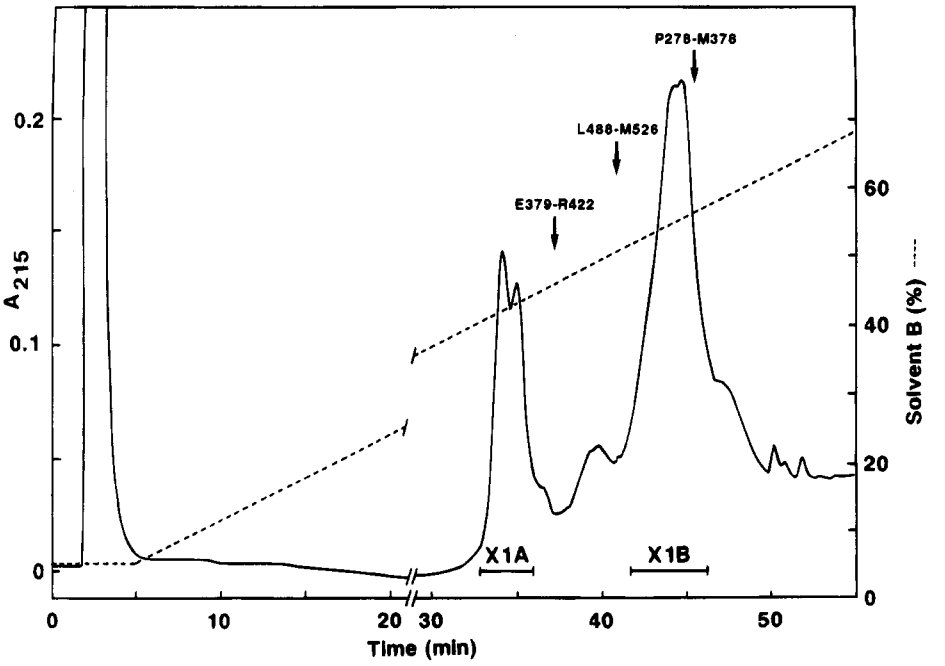

FIGURE 4: Analysis by reverse-phase HPLC of the peptides contained in fraction X1. The peptides contained in fraction X1 (see Figure 2) were analyzed by reverse-phase HPLC on a C4 column, as described under Experimental Procedures. Peptides were detected by absorbance at 215 nm, and pools were made as indicated by bars. The arrows indicate the elution positions of peptides E379−R422, L488−M526, and P278−M378 obtained from CNBr cleavage of the non-cross-linked γ-B sample.

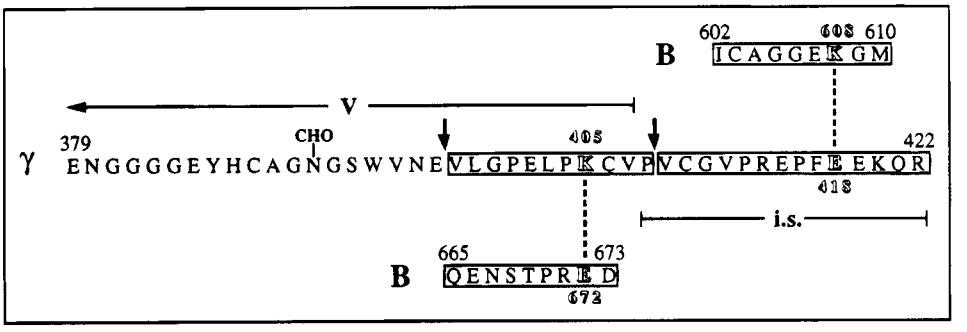

FIGURE 5: Proposed structure of the cross-linked peptides isolated from fraction X1A. Peptides I602−M610 and Q665−D673 both originate from the B chain, and peptide E379−R422 is from the γ fragment. The portions of peptide E379−R422 corresponding to module V and the intermediary segment are indicated. The major sites of cleavage by thermolysin are shown by arrows. The cross-linked peptides isolated after thermolytic cleavage are boxed, and the proposed cross-linking sites are represented by dotted lines. CHO indicates the N-linked oligosaccharide.

Identification of the Cross-Linking Sites. The presence of three peptides in fraction X1A suggested the occurrence of at least two distinct cross-linking sites. With the view to generate short peptides containing only one of these sites, the peptide material contained in fraction X1A was submitted to thermolytic cleavage, as described under Experimental Procedures. The digest was then fractionated by reverse-phase HPLC (data not shown), yielding six major fractions, the peptide contents of which were determined by both N-terminal sequence and mass spectrometry analyses. The cleavage reaction was not complete, as shown by the presence of residual uncleaved material, eluting in the latest fractions, 5 and 6. Both fractions 1 and 2 contained non-cross-linked peptides: the major component of the heterogeneous fraction 1 was peptide E379−E397, originating from the N-terminal end of peptide E379−R422, and fraction 2 contained peptide V409−R422, derived from the C-terminal end of this same peptide (Figure 5).

In contrast, N-terminal sequence analysis of fraction 3 yielded two sequences, corresponding to peptides I602−M610 and V409−R422. Quantitative analysis of the sequence data (Figure 6A) revealed that the PTH derivative of lysine, expected at cycle 7 of the sequence of peptide I602−M610, was not detected, clearly indicating that this residue was involved in the cross-link. Consistent with this conclusion, the only other lysine residue present in the two sequences, occurring at position 420 of peptide V409−R422, was recovered with the expected yield. Of the three acidic amino acids present in peptide V409−R422 (Glu 415, Glu 418, Glu 419), Glu 418 was recovered with the lowest yield (Figure 6A). The decrease was slight but significant, as it was observed consistently upon repeated sequence analyses performed on peptide material obtained from two separate experiments. Sequence analysis of the free form of peptide V409−R422, isolated from fraction 2 (Figure 6B), allowed direct comparison of the yields of Glu 418 in both cases, indicating that the relative decrease was indeed 40−50%. Analysis of the material contained in fraction 3 by both FAB and electrospray ionization mass spectrometry gave evidence for the presence of a species of molecular mass 2701.5 ± 1.0 (about one-third of the total peptide material), corresponding to peptides I602−M610 and V409−R422 cross-

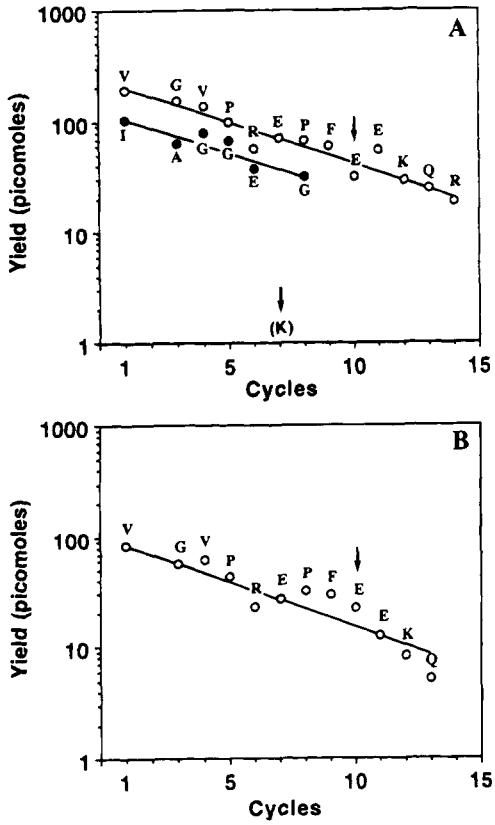

FIGURE 6: Quantitative analysis of the Edman degradation of the cross-linked peptides V409–R422/I602–M610. (A) Sequence analysis of the cross-linked material isolated from the thermolytic cleavage fraction 3, showing the sequences of both peptides V409–R422 (○) and I602–M610 (●). (B) Comparative sequence analysis of the free form of peptide V409–R422, isolated from the thermolytic cleavage fraction 2. The PTH derivatives of *S*-(pyridylethyl)cysteine, expected at cycle 2 of both peptides V409–R422 and I602–M610, and of homoserine, expected at cycle 9 of peptide I602–M610, were effectively identified but not quantified. Arrows indicate residues Lys 608 and Glu 418 involved in the cross-link.

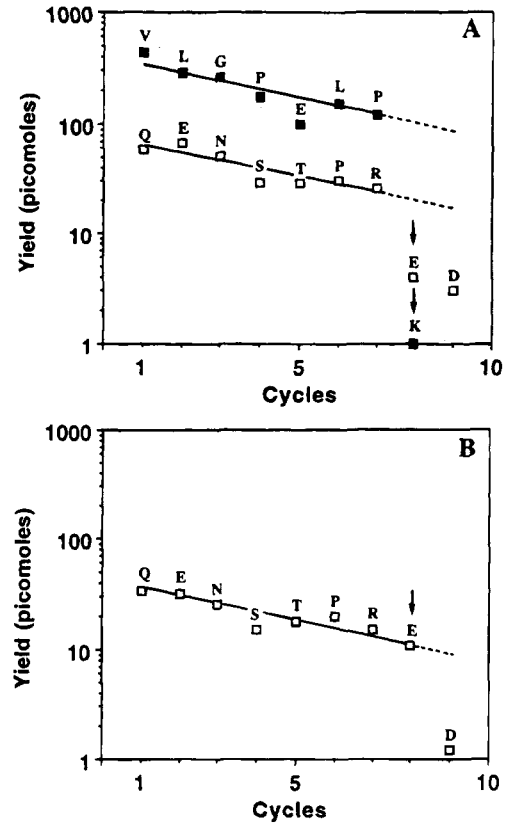

FIGURE 7: Quantitative analysis of the Edman degradation of the cross-linked peptides V398–P408/Q665–D673. (A) Sequence analysis of the cross-linked material isolated from fraction 4 of Figures 5, showing the sequences of both peptides V398–P408 (■) and Q665–D673 (□). (B) Comparative sequence analysis of the synthetic peptide Q665–D673. The PTH derivative of *S*-(pyridylethyl)cysteine, expected at cycle 9 of peptide V398–P408, was effectively identified but not quantified. Arrows indicate residues Lys 405 and Glu 672 involved in the cross-link.

linked through a pseudopeptide bond (expected mass of the species containing an homoserine at position 610 = 2701.2). Two other species corresponding to the free form of peptide V409–R422, either unmodified (observed mass = 1778.5 ± 0.5; expected mass = 1779.1) or EDC-modified (observed mass = 1935.1 ± 0.2; expected mass = 1935.1), each representing about one-third of the total peptide material, were also observed. Consistent with this result, the sequence analysis shown in Figure 6A also indicated the presence, beside the cross-linked species, of the free form of peptide V409–R422, estimated at about 50% of the peptide material. It became clear that one-half to two-thirds of the material submitted to sequence analysis consisted of the free form of peptide V409–R422, providing an explanation for the relative abundance of Glu 418. It was concluded, therefore, that peptides I602–M610 and V409–R422 were cross-linked between the ϵ-amino group of Lys 608 and the γ-carboxyl group of Glu 418, as illustrated in Figure 5.

Again, Edman degradation of the following thermolytic cleavage fraction 4 yielded two sequences, corresponding to peptides V398–P408 and Q665–D673 (see Figure 5). Quantitative analysis of the sequence data (Figure 7A) showed that the overall yield of peptide V398–P408 was much higher than that of peptide Q665–D673, suggesting partial cyclization of the N-terminal glutamine 665 to pyroglutamic acid. Only trace amounts of the PTH derivative of lysine were detected at cycle 8 of peptide V398–P408, indicating that Lys 405, the only lysine residue present in the two sequences, was involved in the cross-link. In the case of peptide Q665–D673, two acidic amino acids, Glu 672 and the C-terminal Asp 673, were recovered with low yields. With a view to discriminate between these residues, sequence analysis was performed on an equivalent amount of the corresponding synthetic peptide (Figure 7B). Comparison of the two analyses indicated that the yield of Glu 672 was decreased by about 85% in the cross-linked sample, whereas that of Asp 673 was equivalent in both cases. It was concluded, therefore, that peptides V398–P408 and Q665–D673 were linked to each other through a pseudopeptide bond involving Lys 405 and Glu 672. Further analysis of fraction 4 by electrospray ionization mass spectrometry clearly showed that peptides V398–P408 and Q665–D673 were cross-linked. Indeed, as summarized in Table 1, several species were detected, corresponding to the unmodified cross-linked species and to modified forms arising from either modification with EDC or cyclization of glutamine 665 to pyroglutamic acid or from both modifications. Significant amounts of the unmodified free form of peptide V398–P408 (observed mass = 1256.8 ± 0.1; expected mass = 1256.6) were also observed, indicating that, as found above for the other site, cross-linking was incomplete.

Homology Modeling of the Protein Modules of the C1s γ-B Region. Bovine chymotrypsin was chosen as the primary template for modeling the catalytic B chain of C1s.

Table 1: Identification by Electrospray Ionization Mass Spectrometry of the Cross-Linked Peptides V398−P408/Q665−D673

species	exptl mass	calcd mass
cross-linked peptides V398−P408/Q665−D673	2313.2 ± 0.2	2313.6
cross-linked peptides V398−P408/Q665−D673 + 1 EDC	2469.6 ± 0.5	2469.6
cross-linked peptides V398−P408/Q665−D673 with pyroglutamic acid	2296.4 ± 0.5	2296.6
cross-linked peptides V398−P408/Q665−D673 with pyroglutamic acid + 1 EDC	2452.5 ± 0.8	2452.6

The sequence alignment used as a guideline was similar to that defined by Greer (1990). The two proteases show 33% sequence identity, and as much as 74% of their residues were considered to be structurally conserved, i.e., superimposed within the limit of 1.5 Å. Compared with chymotrypsin, the C$\bar{1}$s B chain exhibits three deletions at positions 441−442 (four residues), 560−561 (two residues), and 645−646 (three residues) and six minor insertions (Arg 466, Arg 481, Thr 538−Ser 539, Thr 598, Lys 608, Asn 627−Lys 629). There are three major modifications consisting of seven- and eight-residue insertions (Val 504−Asn 510 and Val 583−Ala 590) occurring in two loops, and a five-residue extension (Thr 669−Asp 673) at the C-terminal end of the chain. As expected, the core of the protease, containing the catalytic triad, is highly conserved, whereas all of the modifications are located on the surface. Thus, the two extended loops (Val 504−Asn 510 and Val 583−Ala 590) are located on one side of the entrance of the active site cleft (see Figure 9). Interestingly, the three deletions all occur within the same area, also in the vicinity of the active site cleft entrance. It should be noticed that the C-terminal extension (Thr 669−Asp 673) was constructed as the continuity of the C-terminal α-helix of chymotrypsin (see Figure 9). These three major modifications are the less constrained and therefore the less reliable parts of the model, due to the lack of good templates. From a functional point of view, it should be emphasized that the model is consistent with a classical activation mechanism involving formation of a salt bridge between the N-terminal Ile 423 and Asp 616 and that the S1 subsite (Asp 611), giving trypsin-like specificity, is located in a well-conserved region.

The 15-residue intermediary segment (Pro 408−Arg 422) which is C-terminal to the γ region in C$\bar{1}$s, and links CCP module V and the B chain moiety in proenzyme C1s (Figure 3), was constructed by using the activation peptide of bovine chymotrypsinogen, also comprising 15 residues, as a template structure. The reason for this choice is that the C-terminal end of this segment is structurally better defined in chymotrypsinogen than in chymotrypsin. As in chymotrypsinogen, this segment is covalently linked in C1s to the serine protease domain through a disulfide bridge involving Cys 410 and Cys 534 (Figure 3). As described below, this disulfide bridge is one of the constraints which were used to fit the intermediary segment to the B chain. In this respect, it should be mentioned that Cys 534 of C1s and its counterpart in chymotrypsinogen (Cys 122) are found to be exactly superimposed. In contrast, there is a local displacement between Cys 410 of C1s and the corresponding residue (Cys 1) of chymotrypsinogen, due to their respective location at positions 3 and 1 in the two homologous segments. However, the Cys 410−Cys 534 Cα−Cα distance in the C1s model appears to be compatible with the formation of a disulfide bridge.

A problem for modeling the CCP modules IV and V of C1s arose from the restricted number of homologous structures available, namely, those of the 5th, 15th, and 16th modules from human factor H (Norman et al., 1991; Barlow et al., 1992, 1993). The sequences of these three modules were aligned with the aid of the hydrophobic cluster analysis (HCA) method (Gaboriaud et al., 1987), on the basis of the amino acids considered to form the hydrophobic core, as defined from the three-dimensional structure of modules H5 and H16 (Norman et al., 1991; Barlow et al., 1992). Modeling of modules IV and V of C1s was realized on the basis of the structures of modules H16 and H5, respectively, using the alignments shown in Figure 8. The overall sequence homology (including identities and conservative replacements) is 32% for the C1sIV/H16 pair and 29% for the C1sV/H5 pair. In contrast, if one only takes into account the core residues, as defined in Figure 8, the homology increases to 57% and 77%, respectively, clearly indicating that both pairs of modules have a common framework. Indeed, the resulting three-dimensional models of modules IV and V of C1s both show a five-β-stranded globular structure typical of the CCP modules, with some modifications, located in the C-terminal region of the modules and mostly occurring at the surface. Thus, compared to H16, C1sIV exhibits a single one-residue deletion, two minor insertions (Asn 329, Asn 334), and a six-residue insertion (Val 313−Gly 318) forming a bulge within the third β-strand (Figure 8). In the same way, compared to H5, C1sV exhibits a one-residue deletion, three minor insertions (Glu 348, Gly 381−Gly 382, and Asn 391, which is the N-glycosylation site), and a six-residue insertion (Val 398−Leu 403), extending a preexisting bulge within the fifth β-strand (Figures 8 and 9).

Assembly of the C-Terminal Part of the C$\bar{1}$s γ-B Region. Since cross-linking experiments provided no evidence for an interaction between the N-terminal module IV and the remainder of the γ-B region, assembly was restricted to module V, the intermediary segment, and the B chain. As mentioned above, the intermediary segment (Pro 408−Arg 422) was initially fitted to the surface of the B chain by means of the Cys 410−Cys 534 disulfide bridge, using the same orientation as that of the activation peptide in chymotrypsinogen. In a second step, module V was connected to the intermediary segment through the Val 407−Pro 408 peptide bond. At this stage of the assembly process, the three peptides involved in the cross-linking (E379−R422, I602−M610, Q665−D673) were known, but the cross-linking sites had not been identified. Nevertheless, it was immediately obvious that segments I602−M610 and Q665−D673 were located at two opposite faces of the B chain, each in the vicinity of one of the extremities of segment E379−R422. Indeed, there was a strong indication that one of the cross-links was between I602−M610 (B chain) and the C-terminal end of E379−R422 (intermediary segment). As a consequence, the other cross-link was very likely between Q665−D673 (C-terminal extension of the B chain) and the N-terminal part of E379−R422 (module V). Identification of the cross-linking sites confirmed these hypotheses. Thus, the cross-link between Glu 418 and Lys 608 was consistent with the model (Figure 9), allowing a more precise orientation of the C-terminal end of the intermediary

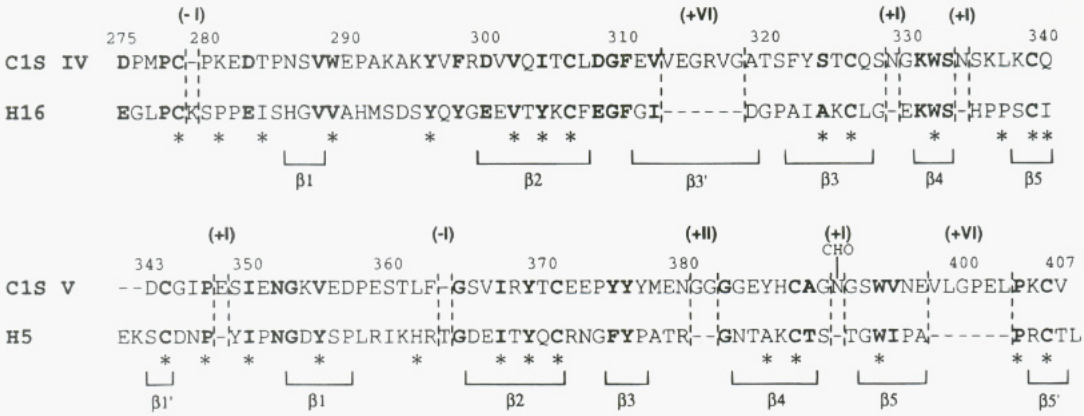

FIGURE 8: Sequence alignments used for modeling the CCP modules IV and V of C1s. The amino acid numbering used is that of C1s. Identical and conservatively replaced residues are shown in bold letters. $\beta1-\beta5$ are the β-strands of the 16th and 5th CCP modules of factor H, and residues considered to form the hydrophobic core of these modules are indicated by stars (Norman et al., 1991; Barlow et al., 1992). Insertions and deletions found in the CCP modules of C1s are shown by Roman numerals in parentheses. CHO indicates the N-linked oligosaccharide.

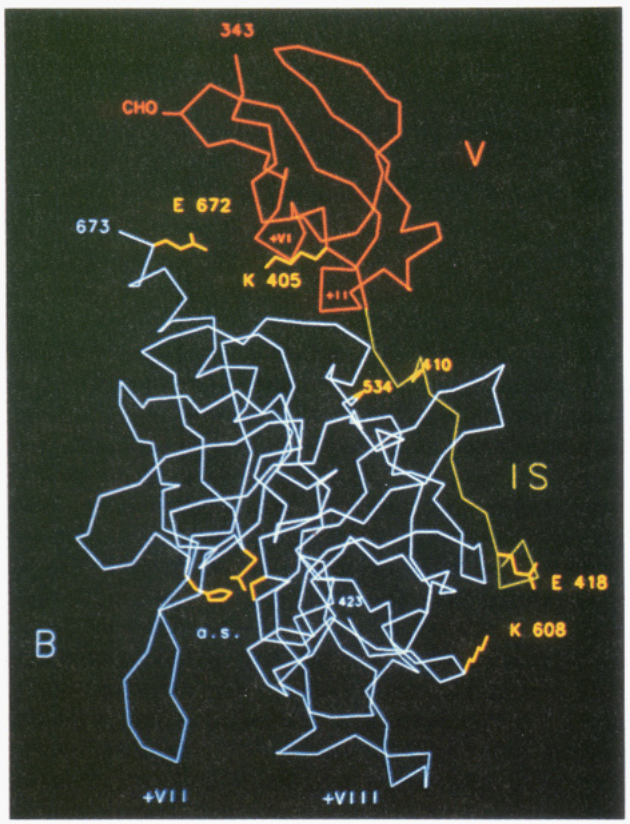

FIGURE 9: Three-dimensional Cα model of the C-terminal part of the catalytic region of C1s. The CCP module V, the intermediary segment (IS), and the B chain are shown in orange, green, and blue, respectively. The side chains of the amino acids involved in the cross-links and of the active site amino acids (His 460, Asp 514, Ser 617), as well as the α carbons of the half-cystine residues Cys 410 and Cys 534, are shown in yellow. The major insertions of module V (Val 398−Leu 403 = +VI; Gly 381−Gly 382 = +II) and of the B chain (Val 504−Asn 510 = +VII; Val 583−Ala 590 = +VIII) are indicated. Abbreviations: CHO, N-linked oligosaccharide; a.s., active site.

segment, a region known to be flexible in active serine proteases. Identification of the other cross-link between Lys 405 and Glu 672 provided a distance constraint for the relative positioning of module V and the C-terminal extension of the B chain. In this case, some uncertainty arises from the fact that, contrary to Lys 405, located in a conserved area of module V (Figure 8), Glu 672 lies in the C-terminal extension of the B chain, arbitrarily modeled as the continuity of the preceding α-helix. Despite their relative inaccuracy, the constraints brought about by the two cross-links and the Cys 410−Cys 534 disulfide bridge allow the relative positioning of module V, the intermediary segment, and the catalytic B chain. They also provide evidence that module V interacts with the B chain on the side opposite to the active site cleft (Figure 9).

DISCUSSION

Chemical cross-linking with EDC is a frequently used method to identify ionic bonds involved in the interaction between two proteins or between individual domains within a protein. An advantage of this zero-length cross-linking agent is that any accessible carboxyl group can be activated and that, provided that it is close to the ϵ-amino group of a lysine residue, a cross-link may occur, thereby indicating that these groups form a salt bridge in the native protein(s). However, hydrophobic interactions, or even ionic interactions involving arginine and aspartic or glutamic acid, are not detected by this technique. The experiments reported in the present study also indicate that "abortive" reactions of EDC with carboxyl groups may occur, as shown by the identification, using mass spectrometry, of EDC-modified forms of peptides P278−M378 and V409−R422 and of the cross-linked peptides V398−P408/Q665−D673. Such reactions complicate analysis of the data, particularly peptide mapping and mass spectrometry analysis, and may, in some cases, interfere with identification of the cross-linked amino acids by N-terminal sequence analysis. Unambiguous identification of the amino acids involved in each cross-link could nevertheless be achieved from comparative sequence analysis of the corresponding unmodified peptide. Despite the limitations of the cross-linking technique, it should be emphasized that the experiments reported in this study were performed twice and that identical results were obtained in each case. As judged from SDS−PAGE analysis of the cross-linking between the γ fragment and the B chain, the overall yield of the reaction was 50−60%. It should be considered, however, that either of the two cross-links identified in this study (Lys 405−Glu 672 and Glu 418−Lys 608) is expected to yield a cross-linked γ-B molecule. It is likely, therefore, that the cross-linked γ-B species isolated by high-pressure gel permeation and hence the

material contained in pool X1A each represent a mixture of species containing either one or the other cross-link, or both of them. Indeed, further analysis of the two cross-links after thermolytic digestion gives evidence, in both cases, for the presence of non-cross-linked peptides, indicating that both cross-linking reactions were incomplete.

Three-dimensional homology modeling of the serine protease domain of C$\bar{1}$s was realized according to the method defined by Greer (1990), using bovine chymotrypsin as a primary template. Although the resulting model is comparable to that obtained previously by Carter et al. (1984), there are noticeable differences with respect to the lengths and the precise boundaries of the major insertions, in particular those corresponding to the two extended loops located in the vicinity of the entrance of the active site cleft. The main reason for these discrepancies probably lies in the fact that the alignment used in the present study is based on a set of reference structures, whereas that used by Carter et al. was based on chymotrypsin alone. Despite the restricted number of reference structures available for CCP modules, the models obtained for modules IV and V of C1s, based respectively on the coordinates of the 16th and 5th modules of factor H, may be considered satisfactory, owing to the particularly high homology shown by each of these pairs of modules. Indeed, compared with the template structures, the models derived for modules IV and V each exhibit a single major insertion and likely provide, therefore, a reliable basis for the major part of their structure.

The information provided by both the chemical cross-linking experiments and the modeling studies allowed us to construct a model of the C-terminal part of the C$\bar{1}$s γ-B region, comprising module V, the 15-residue intermediary segment, and the serine protease B chain. In this regard, it should be emphasized that both approaches not only were complementary but also were consistent with each other. The three protein elements of the assembly are held together by means of three major constraints, of which two (the cross-link between Lys 405 and Glu 672 and the Cys 410−Cys 534 disulfide bridge) allow the relative positioning of module V and the B chain, indicating that these domains are likely closely associated to each other. Indeed, other interactions, not detectable by the cross-linking technique, could take place. For example, additional contacts may be provided by the six- and two-residue insertions of module V which, according to our model, form two bends possibly protruding toward the B domain (Figure 9). The occurrence of a tight association between module V and the B chain would be consistent with the neutron scattering studies performed on the γ-B fragment of C$\bar{1}$s (Zaccaï et al., 1990). It is also compatible with the calorimetric studies reported by Medved et al. (1989), indicating that modules IV and V of C1s unfold independently, at about the same temperature (60 °C), both in the intact protein and in the isolated γ fragment. Indeed, if one takes into account that the B chain melts at a lower temperature (49 °C), the proposed interaction between module V and the B chain cannot be expected to modify extensively the thermal stability of module V. This association between module V and the B chain represents the first example of an interaction between a CCP module and a serine protease domain. It is very likely also a characteristic of human C$\bar{1}$r, given the fact that this protease homologous to C$\bar{1}$s also contains the disulfide bridge (Cys 434−Cys 560) connecting the intermediary segment to the B chain and that EDC treatment yields evidence for a cross-link between Lys 426 of module V and the C-terminal Asp 688 of the B chain, homologous to the Lys 405−Glu 672 cross-link identified in C$\bar{1}$s (M. B. Lacroix, J. Ulrich, N. Scherrer, J. Gagnon, N. M. Thielens, and G. J. Arlaud, unpublished experiments).

The fact that module V of C$\bar{1}$s interacts with the B chain on the side opposite to both the activation site and the catalytic site (Figure 9) allows unimpeded access of the active site of C$\bar{1}$r to the Arg 422−Ile 423 bond cleaved upon activation, as well as of the active site of C$\bar{1}$s to its substrates C4 and C2. With respect to the precise functional role of module V, a likely hypothesis is that it participates in the recognition of one or both of these relatively large substrates (respective M_r 205 000 and 102 000) and allows their correct positioning relative to the active site. Indeed, previous studies based on the use of monoclonal antibodies (Matsumoto & Nagaki, 1986; Matsumoto et al., 1989) give evidence for the presence of a C4 binding site (presumably recognizing the C4b moiety) within the γ region of C$\bar{1}$s. Our model strongly suggests that this site is located within module V and provides a structural basis to test this hypothesis, e.g., by measuring the ability of peptides corresponding to selected portions of module V to inhibit C4 cleavage by C$\bar{1}$s. Other studies (Bing et al., 1978) were based on affinity labeling of C$\bar{1}$s by *m*-[[*o*-(2-chloro-5-(fluorosulfonyl)phenylureido)-phenoxy]butoxy]benzamidine, a reagent designed to bind reversibly through its benzamidine moiety to the S1 subsite (Asp 611) and to form covalent bonds, through its sulfonyl fluoride group, with residues adjacent to the binding site. As the label was found to be distributed on both the A and B chains, it was concluded that portions of the A chain are close to the active site. If one considers the length of the reagent, as calculated by the authors (17 Å), and the distance between Asp 611 and the proximal residues of module V, as estimated from our model (31 Å), it is highly unlikely that the observed labeling occurred in module V. In fact, according to the model shown in Figure 9, the only plausible hypothesis would be a reaction with the closest portion of the intermediary segment.

It is noteworthy that the cross-linking experiments performed in this study yielded no evidence for an interaction between CCP module IV and any of the other domains of the γ-B region. For the reasons mentioned above, this does not prove that there is no such interaction. However, this observation should be considered in light of the NMR structural studies performed on a pair of CCP modules (the 15th and 16th modules of human factor H), indicating the presence of a flexible "hinge" allowing a certain degree of freedom in the relative orientation of the modules (Barlow et al., 1993). If this feature also applies to C$\bar{1}$s, then module IV is expected to adopt a variety of orientations with respect to the remainder of the γ-B region, and therefore it is unlikely that cross-links will be detected. According to current models of the C1 complex (Schumaker et al., 1986; Weiss et al., 1986; Arlaud et al., 1987), activation of the catalytic region of C1s requires a contact with the corresponding regions of C$\bar{1}$r, located **inside** the complex, whereas, once activated, these same regions of C1s are expected to move **outside** the complex, in order to gain access to their substrates C4 and C2. It is tempting to speculate that the occurrence of a flexible hinge between modules IV and V of C1s is responsible for this movement.

ACKNOWLEDGMENT

We are indebted to Prof. I. D. Campbell and Dr. P. N. Barlow, who kindly sent us the coordinates of the 5th, 15th, and 16th CCP modules of human factor H. We thank Y. Petillot for mass spectrometry analysis and I. Bally for peptide synthesis.

REFERENCES

Abola, E. E., Bernstein, F. C., Bryant, S. H., Koetzle, T. F., & Weng, J. (1987) in *Crystallographic Database—Information Content, Software Systems, Scientific Applications* (Allen, F. H., Bergerhoff, G., & Sievers, R., Eds.) pp 107−132, Data Commission of the International Union of Crystallography, Bonn/Cambridge/Chester.

Arlaud, G. J., Sim, R. B., Duplaa, A. M., & Colomb, M. G. (1979) *Mol. Immunol. 16*, 445−450.

Arlaud, G. J., Colomb, M. G., & Gagnon, J. (1987) *Immunol. Today 8*, 106−111.

Barany, G., & Merrifield, R. B. (1980) in *The Peptides* (Gross, E., & Meienhoffer, J., Eds.) Vol. 2, pp 1−284, Academic Press, New York.

Barlow, P. N., Steinkasserer, A., Horne, T. J.; Pearce, J., Driscoll, P. C., Sim, R. B., & Campbell, I. D. (1992) *Biochemistry 31*, 3626−3634.

Barlow, P. N., Steinkasserer, A., Norman, D. G., Kieffer, B., Wiles, A. P., Sim, R. B., & Campbell, I. D. (1993) *J. Mol. Biol. 232*, 268−284.

Bartunik, H. D., Summers, L. J., & Bartsch, H. H. (1989) *J. Mol. Biol. 210*, 813−828.

Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Jr., Brice, M. D., Rodgers, J. R., Kennard, O., Shomanouchi, T., & Tasumi, M. (1977) *J. Mol. Biol. 112*, 535−542.

Bing, D. H., Laura, R., Andrews, J. M., & Cory, M. (1978) *Biochemistry 17*, 5713−5718.

Birktoft, J. J., & Blow, D. M. (1972) *J. Mol. Biol. 68*, 187−240.

Bode, W., Chen, Z., Bartels, K., Kutzbach, C., Schmidt, G., & Bartunik, H. (1983) *J. Mol. Biol. 164*, 237−282.

Carter, P. E., Dunbar, B., & Fothergill, J. E. (1983) *Biochem. J. 215*, 565−571.

Carter, P. E., Dunbar, B., & Fothergill, J. E. (1984) *Philos. Trans. R. Soc. London, B 306*, 293−299.

Cooper, N. R. (1985) *Adv. Immunol. 37*, 151−216.

Fujinaga, M., & James, M. N. G. (1987) *J. Mol. Biol. 195*, 373−396.

Gaboriaud, C., Bissery, V., Benchetrit, T., & Mornon, J. P. (1987) *FEBS Lett. 224*, 149−155.

Greer, J. (1990) *Proteins 7*, 317−334.

Hess, D., Schaller, J., & Rickli, E. E. (1991) *Biochemistry 30*, 2827−2833.

Jones, T. A., & Thirup, S. (1986) *EMBO J. 5*, 819−822.

Jones, T. A., Zou, J. Y., Lowan, S. W., & Kjeldgaard, M. (1991) *Acta Crystallogr., Sect. A 47*, 110−119.

Laemmli, U. K. (1970) *Nature 227*, 680−685.

Mackinnon, C. M., Carter, P. E., Smyth, S. J., Dunbar, B., & Fothergill, J. E. (1987) *Eur. J. Biochem. 169*, 547−553.

Matsudaira, P. (1987) *J. Biol. Chem. 262*, 10035−10038.

Matsumoto, M., & Nagaki, K. (1986) *J. Immunol. 137*, 2907−2912.

Matsumoto, M., Nagaki, K., Kitamura, H., Kuramitsu, S., Nagasawa, S., & Seya, T. (1989) *J. Immunol. 142*, 2743−2750.

Means, G. E., & Feeney, R. E. (1971) in *Chemical Modification of Proteins*, pp 144−148, Holden-Day, Inc., San Francisco.

Medved, L. V., Busby, T. F., & Ingham, K. C. (1989) *Biochemistry 28*, 5408−5414.

Navia, M. A., Mc Keever, B. M., Springer, J. P., Lin, T.-Y., Williams, H. R., Fluder, E. M., Dorn, C. P., & Hoogsteen, K. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 7−11.

Norman, D. G., Barlow, P. N., Baron, M., Day, A. J., Sim, R. B., & Campbell, I. D. (1991) *J. Mol. Biol. 219*, 717−725.

Read, R. J., & James, M. N. G. (1988) *J. Mol. Biol. 200*, 523−551.

Reid, K. B. M., Bentley, D. R., Campbell, R. D., Chung, L. P., Sim, R. B., Kristensen, T., & Tack, B. F. (1986) *Immunol. Today 7*, 230−234.

Remington, S. J., Woodbury, R. G., Reynolds, R. A., Matthews, B. W., & Neurath, H. (1988) *Biochemistry 27*, 8097−8105.

Sawyer, L., Shotton, D. M., Campbell, J. W., Wendell, P. L., Muirhead, H., Watson, H. C., Diamond, R., & Ladner, R. C. (1978) *J. Mol. Biol. 118*, 137−208.

Schumaker, V. N., Hanson, D. C., Kilchherr, E., Phillips, M. L., & Poon, P. H. (1986) *Mol. Immunol. 23*, 556−565.

Schumaker, V. N., Zavodszky, P., & Poon, P. H. (1987) *Annu. Rev. Immunol. 5*, 21−42.

Spycher, S. E., Nick, H., & Rickli, E. E. (1986) *Eur. J. Biochem. 156*, 49−57.

Tam, J. P., Heath, W. F., & Merrifield, R. B. (1983) *Int. J. Pept. Protein Res. 23*, 2939−2942.

Thielens, N. M., Van Dorsselaer, A., Gagnon, J., & Arlaud, G. J. (1990) *Biochemistry 29*, 3570−3578.

Tosi, M., Duponchel, C., Meo, T., & Julier, C. (1987) *Biochemistry 26*, 8516−8524.

Villiers, C. L., Arlaud, G. J., & Colomb, M. G. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 4477−4481.

Wang, D., Bode, W., & Huber, R. (1985) *J. Mol. Biol. 185*, 595−624.

Weiss, V., Fauser, C., & Engel, J. (1986) *J. Mol. Biol. 189*, 573−581.

Zaccaï, G., Aude, C. A., Thielens, N. M., & Arlaud, G. J. (1990) *FEBS Lett. 269*, 19−22.

BI9427781