Human C$\bar{1}$ Inhibitor: Improved Isolation and Preliminary Structural Characterization†

Richard A. Harrison*

ABSTRACT: An improved procedure for the isolation of the C$\bar{1}$ inhibitor (C$\bar{1}$-INH) component of human complement is reported. Following preliminary steps to remove plasminogen, fibrinogen, and aggregated material, three conventional chromatographic steps are used to isolate C$\bar{1}$-INH in high (70%) overall yield. An extinction coefficient ($E^{1\%,1cm}_{280nm}$) of 3.60 has been determined. The isolated protein exhibits a single band on sodium dodecyl sulfate–polyacrylamide gel electrophoresis, with a mobility corresponding to an apparent molecular weight (M_r) of 105 000. After removal of carbohydrate, the protein shows an increased mobility, corresponding to an apparent M_r of 78 000. A total carbohydrate content of 33% has been calculated, and from this and the size of the deglycosylated polypeptide, a true molecular weight of 116 000 was estimated. Further analysis of the carbohydrate has indicated a galactose:mannose ratio of 2:1 and approximately equimolar amounts of *N*-acetylglucosamine and *N*-acetylgalactosamine. This composition is unusual for a plasma protein and suggests that much of the carbohydrate is contained in linkages other than the typical N-glycosidic structures. Values found for the amino acid composition are compared with those reported previously. The amino-terminal sequence (40 residues) of C$\bar{1}$-INH is also reported. Asparagine lies at the amino terminus. Neither high-performance liquid chromatography of the released phenylthiohydantoin derivative nor back-hydrolysis of the thiazolinone permitted identification of the residue contained at position 3. The sequence around this position is compatible, however, with an N-glycosidic linkage to residue 3. The first 10 residues also contain an unusual run of 5 hydroxyl-containing amino acids (-Thr-Ser-Ser-Ser-Ser-) at positions 5–9. Visual comparison of the amino-terminal sequences with those reported for other protease inhibitors does not indicate any sequence homology.

C$\bar{1}$ inhibitor (C$\bar{1}$-INH)[1] is the major homeostatic component of the classical pathway of complement activation, inhibiting the proteolytic activity of activated C1 and thus the activation of C4 and C2. It was first described as a serum α_2-globulin with inhibitory activity against C$\bar{1}$ esterase (Ratnoff & Lepow, 1957; Pensky et al., 1961) and later shown to be identical with the α_2-neuraminoglycoprotein described by Schultze and coworkers (Schultze et al., 1962; Pensky & Schwick, 1969). Analysis of the mechanism by which C$\bar{1}$-INH inhibits C$\bar{1}$ has shown it to act in a stoichiometric fashion with both C$\bar{1}$r and C$\bar{1}$s, forming 1:1 molar complexes (Harpel & Cooper, 1975; Sim et al., 1979b, 1980). The interaction between C$\bar{1}$-INH and both C$\bar{1}$r and C$\bar{1}$s is covalent and believed to involve the active-site serine of the proteases (Arlaud et al., 1979). In this respect, it appears to have a reaction mechanism analogous to those of other plasma protease inhibitors such as α_2-antiplasmin, anti-thrombin III, and α_1-protease inhibitor (Cohen, 1973; Moroi et al., 1975; Owen, 1975; Baugh & Travis, 1976; Moroi & Aoki, 1977; Wiman & Collen, 1979). In addition to this direct inhibitory role, it has recently been shown that C$\bar{1}$-INH promotes dissociation of C$\bar{1}$ bound to immune complexes, releasing a C$\bar{1}$r–C$\bar{1}$s–(C$\bar{1}$-INH)$_2$ complex (Sim et al., 1979a; Ziccardi & Cooper, 1979).

Little is known about the structure of C$\bar{1}$-INH. It is heavily glycosylated, with a reported carbohydrate content of 35% (Haupt et al., 1970), and consists of a single polypeptide chain. Its molecular weight has been estimated as about 105 000 by both $NaDodSO_4$–polyacrylamide gel electrophoresis and analytical ultracentrifugation (Pensky et al., 1961; Haupt et al., 1970; Harpel & Cooper, 1975; Reboul et al., 1977). Recent electron micrographic analysis has suggested that C$\bar{1}$-INH is a highly elongated molecule containing rodlike and globular domains (Odermatt et al., 1981). This fits well with its low sedimentation coefficient of 3.7–4.5 S. Amino acid and carbohydrate compositions of the protein have also been published (Haupt et al., 1970; Harpel et al., 1975).

While C$\bar{1}$-INH was first described as an inhibitor of activated C1, it has long been known that it is active against several other serum proteases. These include components of the coagulation system (factors XIa and XIIa; Forbes et al., 1970), the fibrinolytic system (plasmin; Ratnoff et al., 1969), and the kinin system (kallikrein; Gigli et al., 1970). However, C$\bar{1}$-INH is the only known plasma inhibitor of C$\bar{1}$r and C$\bar{s}$. Other plasma proteases that can be inhibited by C$\bar{1}$-INH have alternative regulatory pathways. For example, plasmin is inhibited by α_2-antiplasmin (Moroi & Aoki, 1976; Collen, 1976; Mullertz & Clemmensen, 1976), and kallikrein can be inhibited by α_2-macroglobulin (Schapira et al., 1981, 1982). It seems probable, therefore, that its major physiological role is directed against complement activation via the classical pathway. These secondary activities, particularly that against kallikrein, remain of interest because of the disease hereditary angioneurotic edema (HANE), which is associated either with low levels of apparently normal C$\bar{1}$-INH (Donaldson & Evans, 1963) or with the presence of a dysfunctional C$\bar{1}$-INH protein (Rosen et al., 1965, 1971).

The work reported here was initiated in order that the interactions between both the normal and dysfunctional C$\bar{1}$-INH

† From the Division of Immunology, Children's Hospital Medical Center and Department of Pediatrics, Harvard Medical School, Boston, Massachusetts 02115. *Received February 18, 1983.* Supported by grants from the Birth Defects Foundation—March of Dimes and Grants AI 05877 and AI 15033 of the U.S. Public Health Service.

* Address correspondence to this author at the MRC Unit on Mechanisms in Tumour Immunity, MRC Centre, Cambridge CB2 2QH, England.

[1] Abbreviations: C$\bar{1}$-INH, C$\bar{1}$ esterase inhibitor [all other complement nomenclature follows the recommendations of the W.H.O. Committee on Complement (1968)]; HANE, hereditary angioneurotic edema; PEG, poly(ethylene glycol); $NaDodSO_4$, sodium dodecyl sulfate; TFMSA, trifluoromethanesulfonic acid; PTH, phenylthiohydantoin; Tris, tris(hydroxymethyl)aminomethane; EDTA, ethylenediaminetetraacetic acid; PAS, periodic acid–Schiff; HPLC, high-performance liquid chromatography.

0006-2960/83/0422-5001$01.50/0

proteins and $C\bar{1}s$ or other plasma proteases, particularly kallikrein, could be investigated. Parallel chemical characterization of the normal and dysfunctional proteins would also permit definition of the structural lesions leading to $C\bar{1}$-INH dysfunction. Plasma containing dysfunctional proteins, though fresh frozen, were frequently several years old (and no longer obtainable). In addition, only limited volumes were available. Considerable attention was therefore spent in developing an isolation procedure, applicable to both normal and dysfunctional proteins, with special emphasis on improving the yield of $C\bar{1}$-INH from long-term-stored plasmas. This, together with preliminary chemical characterization of the isolated functional protein, is reported here.

Materials and Methods

DEAE-Sephadex A50, Sephadex G150 superfine, and Sepharose CL6B were purchased from Pharmacia. Lysine–Sepharose was synthesized according to Deutsch & Mertz (1970). Hydroxylapatite (Bio-Gel HTP) was from Bio-Rad, as were $NaDodSO_4$, acrylamide, methylenebis(acrylamide), *N,N,N′,N′*-tetramethylethylenediamine, Coomassie brilliant blue R250, and Basic Fuchsin. Dithiothreitol was from Calbiochem and iodoacetic acid from Kodak. Iodo[1-^{14}C]-acetic acid was from New England Nuclear. All sequenator reagents and solvents as well as methanol, norleucine, and PTH-norleucine were purchased from Pierce. Acetyl chloride was purchased from Sequemat, and acetonitrile and dichloroethane were from Burdick and Jackson. Anhydrous TFMSA was supplied by Sigma and anisole by Aldrich. Monospecific goat anti-human $C\bar{1}$-INH was provided by Atlantic Antibodies. Fresh frozen plasma was obtained from the American Red Cross Blood Services, Northeast Region. Occasionally, freshly drawn plasma from fasted volunteers was used.

$C\bar{1}$-INH was located in column effluents by single-dimension crossed immunoelectrophoresis (Laurell, 1966) and quantitated antigenically by double diffusion in agar (Mancini et al., 1965). Functional assays were performed by using the procedure of Gigli and co-workers (Gigli et al., 1968). Protein solutions were concentrated by ultrafiltration using PM10 membranes (Amicon). Phosphate buffers were made by using KH_2PO_4 and adjusting the pH with NaOH. $NaDodSO_4$–polyacrylamide gel electrophoresis was performed according to Laemmli (1970). Samples were incubated at 100 °C for 2 min in sample buffer containing 0.1 M 2-mercaptoethanol prior to loading and the gels stained either with Coomassie brilliant blue R250 or by the periodic acid–Schiff (PAS) procedure (Segrest & Jackson, 1972).

Carbohydrate was removed from $C\bar{1}$-INH by incubation in anhydrous TFMSA (Edge et al., 1981). Briefly, the procedure used was as follows: 0.5 mL of anhydrous TFMSA containing 5% anisole was added to 10 mg of salt-free lyophilized protein in a conical stoppered glass tube; the tube was flushed with nitrogen, stoppered, and sealed with Parafilm. The dissolved protein was incubated at 0 °C for 3 h (during which time a faint pint color developed) and then the reaction terminated by the gradual addition, with vortexing, of anhydrous sodium carbonate to neutralize the acid. The neutralized reaction was transferred with excess water to a dialysis bag (Spectrapor 2); excess salt, reactants, and released sugars were removed by dialysis against distilled water, and the precipitated polypeptide was recovered by lyophilization. Alkylation was performed by using the following procedure. Lyophilized proteins (10 mg) was dissolved in 1 mL of 0.2 M Tris-HCl/6.0 M guanidine, pH 8.0. Dithiothreitol was then added to 5 mM, and the reaction vessel was flushed with nitrogen, sealed, and incubated at 25 °C for 30 min. The reduced protein was then radioalkylated by the addition of iodo[1-^{14}C]acetic acid (adjusted to a specific activity of about 10 mCi/mmol) at a final concentration of 20 mM. The reaction was carried out (25 °C, 60 min) under nitrogen in the dark and terminated by the addition of 0.1 mL of 2-mercaptoethanol. Excess reactants were removed by dialysis against water, and the protein was recovered by lyophilization.

Samples for amino acid hydrolysis were hydrolyzed (24–96 h; 110°C) under vacuum in 6.0 M HCl (Ultrex) containing 1% phenol and analyzed with a Beckman 121MB analyzer. Samples were prepared for carbohydrate analysis by hydrolysis for 24 h in 1 M methanolic HCl at 85 °C. Mannitol was used as an internal standard. Trimethylsilylation of the liberated sugars [see Bhaskar & Reid (1981)] was performed prior to analysis by gas–liquid chromatography (Clamp et al., 1972). Automated Edman degradation was performed by using a Beckman 890C sequencer modified with a cold trap. Conversion was performed either in 1 M HCl containing 1.0% ethanethiol (Hermondson et al., 1972) or, in those sequencer runs for which the instrument was equipped with a P-6 autoconverter (Sequemat), in methanolic HCl (10 min, 65 °C; acetyl chloride:methanol 1:7 by volume). A 0.1 M Quadrol program (Brauer et al., 1975) was used, and two coupling cycles were performed before the initial cleavage reaction. Phenylthiohydantoin derivatives were identified by high-performance liquid chromatography. A Zorbax ODS column (Du Pont Instruments) equilibrated in 0.01 M sodium acetate, pH 5.5, and developed with acetonitrile gradients modified from those of Zalut and co-workers (Zalut et al., 1980) was used. The identification of certain residues was confirmed by back-hydrolysis in 6 M HCl containing 0.1% stannous chloride (Mendex & Lai, 1975).

The extinction coefficient of $C\bar{1}$-INH was determined in the following way. $C\bar{1}$-INH, at about 4 mg/mL, was dialyzed against distilled water adjusted to pH 7.0 with ammonium hydroxide. The protein was fully soluble under these conditions, and the 280 nm:260 nm extinction ratio was unaltered from that seen in phosphate-buffered saline. The extinction of the protein solution at selected wavelengths was measured and an aliquot of a standard solution of norleucine added. Duplicate aliquots were taken, lyophilized, and hydrolyzed in vacuo for 24 h in 6.0 M HCl as described previously. The protein concentration of the solution was then determined from the hydrolysis data.

Isolation of $C\bar{1}$-INH. The following procedure has been used routinely for plasma volumes of 500–1500 mL.

Step 1: Removal of Aggregated Material, Fibrinogen, and Plasminogen. Freshly drawn human plasma was made 0.01 M in EDTA and benzamidine by the addition of 0.1 volume of a stock solution of 0.1 M EDTA/0.1 M benzamidine, pH 7.0, and chilled to 4 °C. If fresh frozen plasma was to be used, EDTA and benzamidine were added prior to freezing if possible, otherwise immediately on thawing. All subsequent operations were performed at 4 °C. Solid PEG 4000 was then added to the stirred plasma to a final concentration of 5% (w/v) and the plasma stirred for a further 1–2 h. The precipitate, containing aggregated material, factor XIII, and much of the fibrinogen, was removed by centrifugation (10000*g*, 30 min) and discarded.

Step 2: Lysine–Sepharose Chromatography. The PEG 4000 supernatant was applied to a lysine–Sepharose column (200-mL bed volume/1000 mL of plasma) equilibrated with 0.1 M phosphate/0.5 M KCl/0.01 M EDTA/0.005 M benzamidine, pH 7.0, to remove plasminogen. The column was

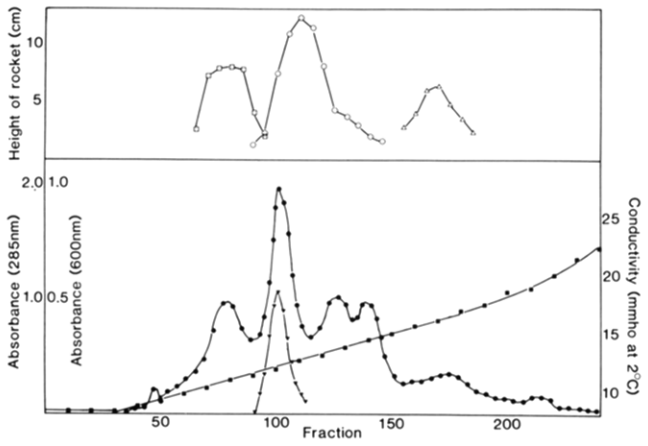

FIGURE 1: Elution profile of the post lysine–Sepharose pool on DEAE-Sephadex A50. Elution conditions as described in the text. (●) A_{285nm}; (▼) A_{600nm}; (■) conductivity; (□) C3; (○) C1̄-INH; (△) C4.

washed with equilibration buffer until the absorbance at 285 nm of the effluent approached zero; all of the unadsorbed protein was pooled.

Step 3: *Chromatography on DEAE-Sephadex A50.* A DEAE-Sephadex A50 column (400-mL bed volume/1000 mL of starting plasma) was equilibrated with 0.02 M phosphate/0.1 M NaCl/0.005 M EDTA/0.005 M benzamidine, pH 7.0. After the pH and conductivity of the lysine–Sepharose effluent were adjusted to those of the equilibration buffer (by dilution and addition of 1.0 M HCl), the pool was applied to DEAE-Sephadex, and after the column was loaded, it was washed with equilibration buffer until the absorbance at 285 nm was below 0.05. The column was then developed with a linear concentration gradient (five column volumes on each side) from 0.02 M phosphate/0.005 M EDTA/0.005 M benzamidine/0.1 M NaCl, pH 7.0, to 0.02 M phosphate/0.005 M EDTA/0.005 M benzamidine/0.4 M NaCl, pH 7.0 (Figure 1).

Step 4: *Chromatography on Sephadex G150 Superfine.* The C1̄-INH-containing fractions from the previous column were pooled, concentrated to between 1 and 2% of the original plasma volume, and dialyzed against 0.02 M phosphate/0.1 M KCl/0.002 M benzamidine/0.002 M citrate, pH 7.0. After dialysis, the protein was clarified by centrifugation (20000*g*, 10 min, 4 °C) and applied to a Sephadex G150 superfine column (2000-mL bed volume/1000 mL of starting plasma) equilibrated with the above buffer. A load volume of 1–2% of the column bed volume was used. A typical elution profile is shown in Figure 2.

Step 5: *Chromatography on Hydroxylapatite.* C1̄-INH-containing fractions from the G150 effluent were located as described previously, pooled, concentrated to between 20 and 30 mL, and dialyzed against 0.02 M phosphate/0.15 M KCl, pH 7.0. The dialyzed protein was then loaded onto a hydroxylapatite column (250-mL bed volume/1000 mL of plasma) equilibrated with the same buffer and elution continued with at least two column volumes of the equilibration buffer. This was followed by a gradient from 0.02 M phosphate/0.15 M KCl, pH 7.0, to 0.02 M phosphate/1.0 M KCl, pH 7.0. Finally, a gradient from 0.02 M phosphate/1.0 M KCl, pH 7.0, to 0.3 M phosphate/1.0 M KCl, pH 7.0, was applied. C1̄-INH-containing fractions (see Figure 3) were pooled, concentrated, and stored frozen at −80 °C.

Results and Discussion

Isolation of C1̄-INH. Because we experienced difficulties, particularly with older plasmas and fibrin precipitation, when

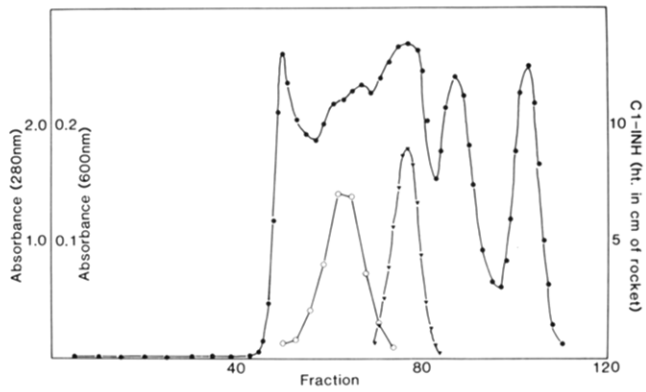

FIGURE 2: Elution profile of the post DEAE-Sephadex A50 pool on Sephadex G150. Elution conditions as described in the text. (●) A_{280nm}; (▼) A_{600nm}; (○) C1̄-INH.

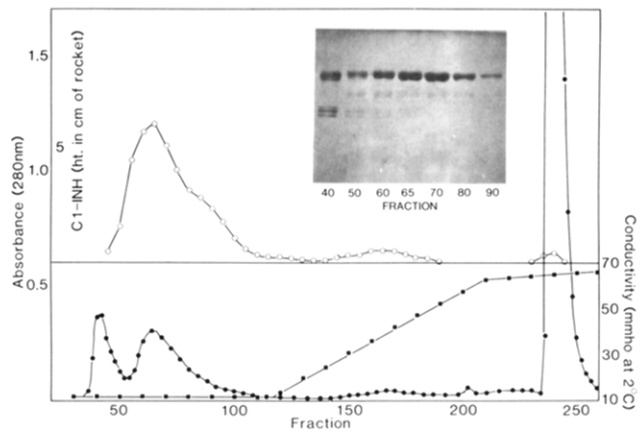

FIGURE 3: Elution profile of the post Sephadex G150 pool on hydroxylapatite. Elution conditions as described in the text. (●) A_{280nm}; (■) conductivity; (○) C1̄-INH. Inset shows $NaDodSO_4$–polyacrylamide gel analysis of fractions eluting before application of the first gradient. Samples were prepared as described under Materials and Methods, and a 10% polyacrylamide gel was used.

using published isolation procedures (Haupt et al., 1970; Harpel & Copper, 1975; Reboul et al., 1977), the procedure described here was developed. The first two steps (5% PEG 4000 precipitation and lysine–Sepharose chromatography), while achieving little in terms of protein fractionation (see Table I), were essential if maximum resolution and recovery were to be made in later chromatographic steps. The advantage of removing plasminogen at an early stage is obvious. It is less clear why PEG precipitation is required as considerable amounts of fibrinogen remain in the supernatant. It is possible that successful isolation of C1̄-INH requires the removal of factor XIII at this point. A relatively high salt concentration (0.5 M KCl) was used during lysine–Sepharose chromatography as at lower salt concentrations considerable non-ligand-specific adsorption was seen.

Resolution of C1̄-INH from the bulk of the plasma proteins was achieved by using three conventional chromatographic steps. Loading conditions onto DEAE-Sephadex were selected such that the albumin was not bound and could therefore be washed out of the column prior to appplication of the gradient. This was important as it was difficult to remove albumin at later stages in the procedure. An additional advantage resulting from these conditions was that C3 and C4 were both retained on the column and eluted during application of the gradient as significantly purified proteins. These proteins could be further purified by using chromatography on Sepharose CL6B and isolated rapidly in high yield and with high specific activity.

Table I: Summary of C$\bar{1}$-INH Isolation

pool	recovery				purification factor
	volume (mL)	A_{280} units	antigen[a]	antigen (%)	
plasma	1254	50 285	1028	100	1
5% poly(ethylene glycol) supernatant	1250	52 750	950	92	0.9
post lysine–Sepharose	1930	48 250	984	96	1.0
post DEAE-Sephadex	820	2 230	1025	99	22.5
post Sephadex G150	108	338	678	66	98.1
post hydroxylapatite	250	59.3	740	72	610

[a] Expressed relative to the amount of C$\bar{1}$-INH contained in 1 mL of a normal plasma pool.

It was important that, with the exception of the final chromatographic step, metal chelators were present throughout the procedure. In their absence, both the early complement proteases C1r and C1s and some coagulation proteases are continuously activated. These can either consume C$\bar{1}$-INH or lead to fibrinogen activation and subsequent precipitation of fibrin. However, if EDTA was included in the elution buffer for Sephadex chromatography, prolonged dialysis was required before application to hydroxylapatite to avoid EDTA-mediated breakdown of the calcium phosphate matrix. Low concentrations of citrate were found to be a suitable alternative for the Sephadex fractionation step. Prolonged elution from hydroxylapatite with benzamidine-containing buffers also caused deterioration of the hydroxylapatite matrix, and benzamidine was therefore omitted from this step.

The equilibration and running conditions for hydroxylapatite chromatography were more critical. For successful fractionation, C$\bar{1}$-INH was slightly retarded under the equilibration conditions and eluted as a broad peak close behind a small breakthrough peak (see Figure 3). Column length was important for resolution of C$\bar{1}$-INH from this initial peak, which contained two proteins of M_r 40 000–45 000 (reduced subunit) as well as a significant amount of an unidentified protein of about M_r 100 000. When short, wide hydroxylapatite columns were used, this protein cochromatographed with C$\bar{1}$-INH, was not resolved by $NaDodSO_4$–polyacrylamide gel electrophoresis, and was therefore an unrecoginized contaminant. However, in order that an acceptable flow rate could be achieved, the column length:width ratio was held at about 20:1. When combined with a small loading volume (about one-tenth column volume), good separation of C$\bar{1}$-INH from unbound proteins was achieved. Occasionally, some C$\bar{1}$-INH did not elute until application of the chloride gradient. This contained appreciable amounts of albumin, had a significantly lower specific activity, and was either pooled separately or discarded. The bulk of the loaded protein was retained on the column and requried higher phosphate concentrations for elution.

Some workers have recommended the use of siliconized glassware or plasticware and irreversible serine protease inhibitors such as diisopropyl phosphorofluoridate or phenylmethanesulfonyl fluoride during early steps of C$\bar{1}$-INH isolation (Sim & Reboul, 1981). We have not found such precautions to be necessary. Recovery of protein and of C$\bar{1}$-INH at each step of a typical isolation is given in Table I, and Figure 4 illustrates the composition of each C$\bar{1}$-INH-containing pool. An overall yield (of antigen) of around 70–75% was routinely found. There was a significant increase in total functional activity during isolation. This probably reflects non-C$\bar{1}$-dependent consumption of C$\bar{1}$-INH during the assay of early pools. Calculation of the plasma concentration of C$\bar{1}$-INH from the overall purification and recovery data together with the calculated extinction coefficient of the isolated protein gives a value of 0.197 mg/mL. This is in close agreement with other published figures (Rosen et al., 1971; Heimburger, 1975).

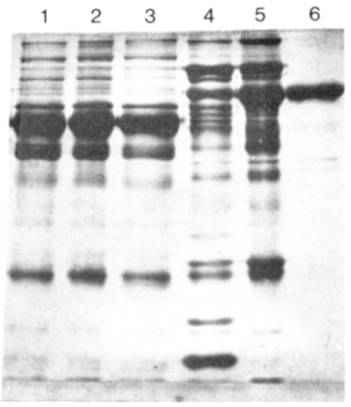

FIGURE 4: $NaDodSO_4$–polyacrylamide gel analysis of the C$\bar{1}$-INH-containing pools at each step of the fractionation. Samples were prepared as described under Materials and Methods, and a 10% polyacrylamide gel was used. Track 1, plasma; track 2, 5% PEG 4000 supernatant; track 3, post lysine–Sepharose pool; track 4, post DEAE-Sephadex A50 pool; track 5, post Sephadex G150 pool; track 6, post hydroxylapatite pool.

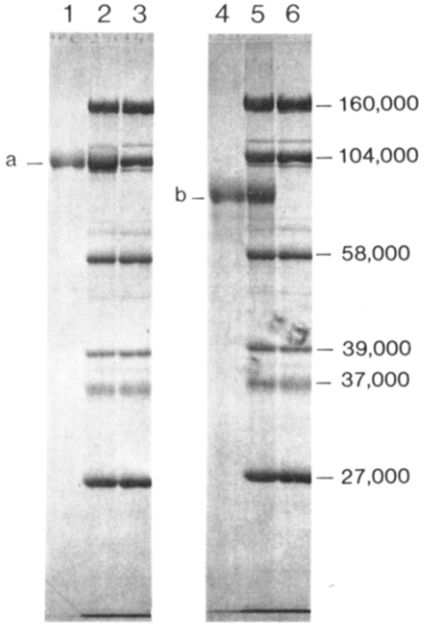

FIGURE 5: Comparison of native C$\bar{1}$-INH with TFMSA-treated C$\bar{1}$-INH. Samples were prepared as described under Materials and Methods and run in the presence of $NaDodSO_4$ on an 8% polyacrylamide gel. Track 1, native C$\bar{1}$-INH (a); track 2, native C$\bar{1}$-INH + marker proteins; tracks 3 and 6, marker proteins; track 4, TFMSA-treated C$\bar{1}$-INH (b); track 5, TFMSA-treated C$\bar{1}$-INH + marker proteins. Marker proteins used were RNA polymerase (β and β', M_r 160 000; α, M_r 39 000), pyruvate decarboxylase (M_r 104 000), catalase (M_r 58 000), D-amino acid oxidase (M_r 37 000), and ribitol dehydrogenase (M_r 27 000).

Physicochemical Characterization of the Isolated Protein. On electrophoresis in $NaDodSO_4$, C$\bar{1}$-INH isolated by using

Table II: Physicochemical Properties of C1-INH

app M_r				105 000	
M_r of carbohydrate-free polypeptide				78 000	
carbohydrate content (% by wt)				33	
true M_r				116 000	
$E^{1\%,1cm}_{280\ nm}$				3.6	
	composition (residues/mol)				
Cys	7.1[a]	Met	11.4	fucose	1.6
Asp	70.4	Ile	19.2	ribose	0.2
Thr	78.3	Leu	85.4	glucose	3.3
Ser	78.3	Tyr	17.8	mannose	21.7
Glu	73.3	Phe	34.2	galactose	41.4
Pro	50.5	Lys	39.1	GlcNAc	33.2
Gly	27.0	His	15.7	GalNAc	29.5
Ala	47.7	Arg	18.5	sialic acid	73.5
Val	38.4	Trp	9.3[b]		

[a] Determined as *S*-(carboxymethyl)cysteine. [b] Calculated from the data of Haupt et al. (1970).

the above procedure gave a single band. This corresponded to an apparent molecular weight of 105 000 (Figure 5). Other workers have reported the existence of a second band of 96 000 apparent molecular weight and have suggested that it arises following proteolytic cleavage of the 105 000 molecular weight polypeptide during isolation (Harpel & Cooper, 1975). This 96 000-dalton polypeptide retained inhibitory activity against C1s and plasmin and was unaltered antigenically. We have seen a similar higher mobility band in stored C1-INH samples but have been unable to generate it reproducibly. Some data suggest that, whether or not proteolytic scission has occurred, there is a considerable alteration in the carbohydrate content of the 96 000-dalton component (R. A. Harrison, unpublished results). Following deglycosylation with trifluoromethanesulfonic acid, the mobility of the polypeptide on NaDodSO$_4$–polyacrylamide gel electrophoresis was increased (Figure 5), corresponding to an apparent molecular weight of 78 000. This band was PAS negative, suggesting that most, if not all, of the carbohydrate had been removed. Since no small peptides were observed, and the amino-terminal sequence and amino acid composition of the protein were unaltered (see Table III), this is believed to reflect the molecular weight of the polypeptide portion of the protein. A carbohydrate content for C1-INH of 33% (by weight) was determined, slightly lower than has previously been estimated (Haupt et al., 1970). This, together with the estimated 78 000-dalton size of the carbohydrate-free polypeptide, allows a true molecular weight of 116 000 to be derived. An $E^{1\%,1cm}_{280nm}$ of 3.6 was found in this study. This is an extremely low value for a plasma protein, presumably reflecting the low tryptophan content of C1-INH, and is 20% lower than that value (4.5) reported in an earlier study (Haupt et al., 1970). These physicochemical properties of the protein, together with its amino acid and carbohydrate composition, are summarized in Table II.

In Table III, the amino acid composition, derived from duplicate 24-, 48-, 72-, and 96-h hydrolyses, is compared with those reported in two earlier studies (Haupt et al., 1970; Harpel et al., 1975). With the exception of the glycine and alanine values, there is reasonable agreement between the three compositions. Further analysis of the carbohydrate contained in C1-INH is contained in Table II. While good agreement with the total hexose content reported by Haupt and coworkers is seen, we have found a galactose:mannose ratio of 2:1 rather than that of 1:2. We also find higher values for hexosamine and sialic acid. Characterization of the hexosamine, not reported before, shows approximately equimolar amounts of *N*-acetylglucosamine and *N*-acetylgalactosamine. The high galactose and *N*-acetylgalactosamine content of

Table III: Comparison of the Amino Acid Composition of Native and Carbohydrate-Free C1-INH[a]

amino acid	*b*	*c*	*d*	*e*
Cys	0.010[f]	*g*	0.010	0.011
Asp	0.099	0.098	0.092	0.088
Thr	0.110	0.105	0.104	0.103
Ser	0.110	0.109	0.089	0.111
Glu	0.103	0.107	0.106	0.087
Pro	0.071	0.070	0.065	0.067
Gly	0.038	0.048	0.029	0.064
Ala	0.067	0.068	0.063	0.093
Val	0.054	0.053	0.065	0.074
Met	0.016	0.007	0.024	0.015
Ile	0.027	0.025	0.036	0.033
Leu	0.120	0.114	0.112	0.107
Tyr	0.025	0.022	0.029	0.011
Phe	0.048	0.049	0.050	0.037
Lys	0.055	0.054	0.062	0.052
His	0.022	0.022	0.021	0.018
Arg	0.026	0.028	0.031	0.016
Trp	*g*	*g*	0.013	*g*

[a] All compositions are expressed as mole fractions of each amino acid. [b] Composition of the native C1-INH protein. Values for serine and threonine are corrected for destruction during hydrolysis. Values for valine and isoleucine represent the maximum amount released with prolonged hydrolysis times. [c] Composition of C1-INH following the removal of carbohydrate with trifluoromethanesulfonic acid. [d] Data from Haupt et al. (1970) expressed as mole fractions. [e] Data from Harpel et al. (1975) expressed as mole fractions. [f] Determined as *S*-(carboxymethyl)cysteine. [g] Not determined.

Chart I: Amino-Terminal Sequence of C1-INH

Asn-Pro-X-Ala-Thr-Ser-Ser-Ser-Ser-Gln-Asp-Pro-Glu-Ser-Leu-Gln-Asp-Arg-Gly-Glu-Gly-Lys-Val-Ala-Thr-X-Val-Ile-Ser-Lys-Met-Leu-Phe-Val-Glu-Pro-Ile-Leu-Glu-Val-

C1-INH is atypical of plasma proteins [see Kornfeld & Kornfeld (1982)]. For example, α_1-protease inhibitor contains two and α_1-acid glycoprotein contains five distinct types of oligosaccharide chains (Fournet et al., 1978; Hodges et al., 1979). These all have approximately equimolar amounts of galactose and mannose, but no *N*-acetylgalactosamine, and are typical complex asparagine-linked units. The high *N*-acetylgalactosamine content of C1-INH, together with the slightly elevated levels of aliphatic hydroxyl-containing amino acids, is in some respects similar to a recently isolated plasma galactoglycoprotein (Schmid et al., 1980) in which the majority of the carbohydrate chains are held via *O*-glycosidic linkages. The unusual carbohydrate composition of C1-INH makes it likely that it too has a number of *O*-glycosidic oligosaccharide units.

Amino-Terminal Sequence Analysis of C1-INH. The amino-terminal sequence of C1-INH (40 residues) is given in Chart I. This sequence was found for duplicate runs on two different preparations of the protein. In addition, the first 12 residues were verified by using the carbohydrate-free (TFMSA-treated) protein. Initial recoveries varied from 20 to 45%, with repetitive yields of 92–96%. Recoveries at each step for a typical run, together with the released counts at each cycle, are given in Figure 6. Several comments should be made about this sequence. At position 3, using both HPLC analysis of the PTH derivative and back-hydrolysis of the thiazolinone, no assignment could be made. Furthermore, monitoring of the released counts did not indicate that there had been specific release of radiolabel (see Figure 6). However, the sequence around this position, with threonine two positions carboxy terminal to the unidentified residue, is compatible with the presence of an N-glycosidic linkage to

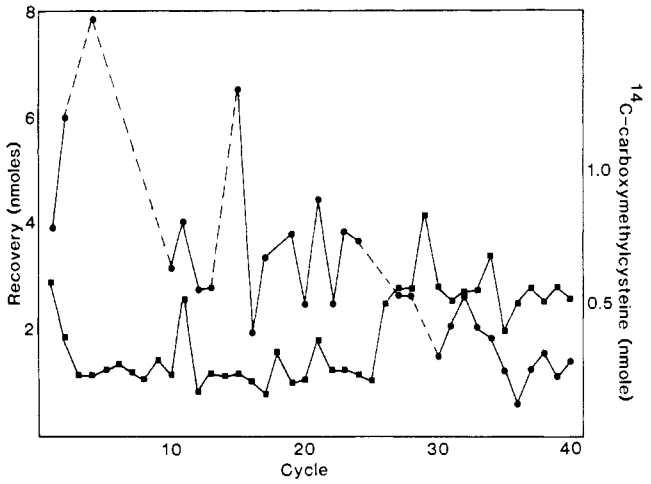

FIGURE 6: Recovery data for the first 40 residues of C1̄-INH from a typical sequencer run. For the run shown, 50 nmol of C1̄-INH was loaded into the cup, and the instrument was fitted with an autoconverter. The repetitive yield, based on recoveries of PTH-Ala at steps 4 and 24 and of PTH-Leu at steps 15, 32, and 38, was 95%. Extrapolation of this recovery for alanine to step 1 gives an initial yield of 30%. The carry-over increased rapidly to 43% at step 4 and seemed related to interference in the cleavage reaction by the residue contained at position 3. From step 4 onward, carry-over increased by only 1% per cycle. If higher initial loads were used, there was an increased carry-over at each cycle. (●) Recovery of PTH-amino acid. (Quantitation of PTH-Arg, PTH-Thr, and PTH-Ser was unreliable, and these figures are not included.) (■) Recovery of ^{14}C, expressed as nanomoles of carboxymethylcysteine.

asparagine (Neuberger et al., 1972). Residues 5–9 comprise an unusual run of hydroxyl-containing amino acids. While recoveries at these positions, particularly of PTH-serine, were low, no other amino acids were detected at levels above background. Confirmation by back-hydrolysis beyond residue 5 was ambiguous because of the high background consequent on washout of uncleaved protein from the cup. Assignment was aided, however, by the identification of two breakdown products of PTH-serine with relative retention times (to PTH-norleucine) of 0.374 and 0.542. The identification of and relative yields of these were confirmed by parallel sequence analysis of the first 20 residues of bovine pancreatic ribonuclease, which contains the sequence -Ser-Ser-Thr-Ser- at positions 15–18 (Smyth et al., 1963). Similar criteria were applied to identification at positions 14 (Ser), 25 (Thr), and 29 (Ser). However, in the absence of independent verification (e.g., by analysis of small amino-terminal peptides), these three assignments should be regarded as tentative. The recoveries from step 10 onwards were consistent with the initial recoveries and repetitive yields. In addition to residue 3, no assignment could be made at position 26.

Other than C1̄-INH, plasma contains six major protease inhibitors [see Laskowski & Kato (1980)]. Of these, α_2-macroglobulin appears unique as it alone possesses an internal thio ester (Tack et al., 1980; Sottrup-Jensen et al., 1980). This enables it to bind covalently to an activating protease, a conformational change consequent on disruption of the thio ester, rendering the active site of the protease inaccessible to macromolecular substrates (Barrett & Starkey, 1973; Barret, 1981; Salvesen et al., 1981). There is no evidence that C1̄-INH utilizes the same reaction mechanism. In contrast, C1̄-INH appears similar to α_1-protease inhibitor, α_2-antiplasmin, and anti-thrombin III in its mechanism of action, forming a covalent complex that involves the active-site serine of the complexed protease. Of these, α_1-protease inhibitor and anti-thrombin III are now known to belong to the same "family", a family which also includes ovalbumin (Carrell et al., 1979, 1982; Hunt & Dayoff, 1980).

Visual comparison of the amino-terminal sequence of C1̄-INH with the sequences of α_1-protease inhibitor, anti-thrombin III, and ovalbumin does not indicate any homology, nor is there any obvious sequence homology between C1̄-INH and other protease inhibitors. However, while no homology was seen between C1̄-INH and the above inhibitor family, this does not necessarily indicate that C1̄-INH is not a member. C1̄-INH is a larger protein, with a polypeptide molecular weight of around 78 000 rather than the values of 41 000–46 000 found for α_1-protease inhibitor, anti-thrombin III, and ovalbumin (McReynolds et al., 1978; Petersen et al., 1979; Carrell et al., 1982). With these proteins, the sequences align at the carboxy-terminal end, with variably extended amino-terminal sections. Thus, a homologous region in C1̄-INH would only extend through approximately half its sequence and might reasonably be expected to lie at the carboxy-terminal rather than the amino-terminal end of the protein. We are now engaged in further analysis of the protein in order to establish its relationship to other protease inhibitors and to comprehend better its mode of action.

Acknowledgments

I thank Semie Awad for invaluable technical assistance, Dr. K. Bhaskar and D. O'Sullivan for performing the carbohydrate analysis, and Professor F. S. Rosen for his constant encouragement.

Registry No. Complement C1̄ inhibitor, 80295-38-1.

References

Arlaud, G. J., Reboul, A., Sim, R. B., & Colomb, M. G. (1979) *Biochim. Biophys. Acta 576*, 151.
Barrett, A. J. (1981) *Methods Enzymol. 80*, 737.
Barrett, A. J., & Starkey, P. M. (1973) *Biochem. J. 133*, 709.
Baugh, R., & Travis, J. (1976) *Biochemistry 15*, 836.
Bhaskar, K. R., & Reid, L. (1981) *J. Biol. Chem. 256*, 7583.
Brauer, A. W., Margolies, M. N., & Haber, E. (1975) *Biochemistry 14*, 3029.
Carrell, R. W., Owen, M. C., Brannan, S. O., & Vaughan, L. (1979) *Biochem. Biophys. Res. Commun. 91*, 1032.
Carrell, R. W., Jeppsson, J. O., Laurell, C. B., Brannan, S. O., Owen, M. C., Vaughan, L., & Boswell, D. R. (1982) *Nature (London) 298*, 329.
Clamp, J. R., Bhatti, T., & Chambers, R. E. (1972) in *Glycoproteins: Their Composition, Structure and Function* (Gottschalk, A., Ed.) 2nd ed., pp 300–321, Elsevier, Amsterdam.
Cohen, A. B. (1973) *J. Biol. Chem. 248*, 7055.
Collen, D. (1976) *Eur. J. Biochem. 69*, 209.
Deutsch, D. G., & Mertz, E. T. (1970) *Science (Washington, D.C.) 170*, 1095.
Donaldson, V. H., & Evans, R. R. (1963) *Am. J. Med. 35*, 37.
Edge, A. S. B., Faltynek, C. R., Hof, L., Reichert, L. E., & Weber, P. (1981) *Anal. Biochem. 118*, 131.
Forbes, C. D., Pensky, J., & Ratnoff, O. D. (1970) *J. Lab. Clin. Med. 76*, 809.
Fournet, B., Montreuil, J., Strecker, G., Dorland, L., Haverkamp, J., Vliegenthart, J. F. G., Binette, J. P., & Schmid, K. (1978) *Biochemistry 17*, 5206.
Gigli, I., Ruddy, S., & Austen, K. F. (1968) *J. Immunol. 100*, 1154.
Gigli, I., Mason, J. W., Colman, R. W., & Austen, K. F. (1970) *J. Immunol. 104*, 574.

Harpel, P. C., & Cooper, N. R. (1975) *J. Clin. Invest. 55*, 593.
Harpel, P. C., Hugli, T. E., & Cooper, N. R. (1975) *J. Clin. Invest. 55*, 606.
Haupt, H., Heimburger, N., Krantz, T., & Schwick, H. G. (1970) *Eur. J. Biochem. 17*, 254.
Heimburger, N. (1975) *Cold Spring Harbor Conf. Cell Proliferation 2*, 367.
Hermondson, M. A., Ericsson, L. H., Titani, K., Neurath, H., & Walsh, K. A. (1972) *Biochemistry 11*, 4493.
Hodges, L. C., Laine, R., & Chan, S. K. (1979) *J. Biol. Chem. 254*, 8208.
Hunt, L. T., & Dayoff, M. O. (1980) *Biochem. Biophys. Res. Commun. 95*, 864.
Kornfeld, R., & Kornfeld, S. (1980) in *The Biochemistry of Glycoproteins and Proteoglycans* (Lennarz, W., Ed.) pp 1–34, Plenum Press, New York.
Laemmli, U. K. (1970) *Nature (London) 227*, 680.
Laskowski, M., & Kato, I. (1980) *Annu. Rev. Biochem. 49*, 593.
Laurell, C. B. (1966) *Anal. Biochem. 15*, 45.
Mancini, G., Carbonara, A. O., & Heremans, J. F. (1965) *Immunochemistry 2*, 235.
McReynolds, L., O'Malley, B. W. O., Nisbet, A. D., Fothergill, J. E., Givol, D., Fields, S., Robertson, M., & Brownlee, G. G. (1978) *Nature (London) 273*, 723.
Mendez, E., & Lai, C. Y. (1975) *Anal. Biochem. 68*, 47.
Moroi, M., & Aoki, N. (1976) *J. Biol. Chem. 251*, 5956.
Moroi, M., & Aoki, N. (1977) *Biochim. Biophys. Acta 482*, 412.
Moroi, M., Yamasaki, M., & Aoki, N. (1975) *J. Biochem. (Tokyo) 78*, 925.
Mullertz, S., & Clemmensen, I. (1976) *Biochem. J. 159*, 545.
Neuberger, A., Gottschalk, A., Marshall, R. D., & Spiro, R. G. (1972) in *Glycoproteins* (Gottschalk, A., Ed.) Vol. 5, Part A, pp 450–490, Academic Press, New York.
Odermatt, E., Berger, H., & Sano, Y. (1981) *FEBS Lett. 131*, 283.
Owen, W. G. (1975) *Biochim. Biophys. Acta 405*, 380.
Pensky, J., & Schwick, H. G. (1969) *Science (Washington, D.C.) 163*, 698.
Pensky, J., Levy, L. R., & Lepow, I. H. (1961) *J. Biol. Chem. 236*, 1674.
Petersen, T. E., Dudek-Wojciechowska, G., Sottrup-Jensen, L., & Magnusson, S. (1979) in *The Physiological Inhibitors of Blood Coagulation and Fibrinolysis* (Collen, D., Ed.) p 43, Elsevier, Amsterdam.
Ratnoff, O. D., & Lepow, I. H. (1957) *J. Exp. Med. 106*, 327.
Ratnoff, O. D., Pensky, J., Ogston, D., & Naff, G. B. (1969) *J. Exp. Med. 129*, 315.
Reboul, A., Arlaud, G. J., Sim, R. B., & Colomb, M. G. (1977) *FEBS Lett. 79*, 45.
Rosen, F. S., Charache, P., Pensky, J., & Donaldson, V. H. (1965) *Science (Washington, D.C.) 148*, 957.
Rosen, F. S., Alper, C. A., Pensky, J., Klemperer, M. R., & Donaldson, V. H. (1971) *J. Clin. Invest. 50*, 2143.
Salvesen, G. S., Sayers, C. A., & Barrett, A. J. (1981) *Biochem. J. 195*, 453.
Schapira, M., Scott, C. F., & Colman, R. W. (1981) *Biochemistry 20*, 2738.
Schapira, M., Scott, C. F., James, A., Silver, L. D., Kueppers, F., James, H. L., & Colman, R. W. (1982) *Biochemistry 21*, 567.
Schmid, K., Mao, S. K. Y., Kimura, A., Hayashi, S., & Binette, J. P. (1980) *J. Biol. Chem. 255*, 3221.
Schultze, H. E., Heide, K., & Haupt, H. (1962) *Naturwissenschaften 49*, 133.
Segrest, J. P., & Jackson, R. L. (1972) *Methods Enzymol. 28*, 54.
Sim, R. B., & Reboul, A. (1981) *Methods Enzymol. 80*, 43.
Sim, R. B., Arlaud, G. J., & Colomb, M. G. (1979a) *Biochem. J. 179*, 449.
Sim, R. B., Reboul, A., Arlaud, G. J., Villiers, C. L., & Colomb, M. G. (1979b) *FEBS Lett. 97*, 111.
Sim, R. B., Arlaud, G. J., & Colomb, M. G. (1980) *Biochim. Biophys. Acta 612*, 433.
Smyth, D. G., Stein, W. H., & Moore, S. (1963) *J. Biol. Chem. 238*, 227.
Sottrup-Jensen, L., Petersen, T. E., & Magnusson, S. (1980) *FEBS Lett. 121*, 275.
Tack, B. F., Harrison, R. A., Janatova, J., Thomas, M. L., & Prahl, J. W. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 5764.
W.H.O. Committee on Complement (1968) *Bull. W.H.O. 39*, 935.
Wiman, B., & Collen, D. (1979) *J. Biol. Chem. 254*, 9291.
Zalut, C., Henzel, W. S., & Harris, H. W. (1980) *J. Biochem. Biophys. Methods 3*, 11.
Ziccardi, R. J., & Cooper, N. R. (1979) *J. Immunol. 123*, 788.