Evidence That Serpin Architecture Intrinsically Supports Papain-like Cysteine Protease Inhibition: Engineering α_1-Antitrypsin To Inhibit Cathepsin Proteases†

James A. Irving,[‡,§,+] Robert N. Pike,[‡,§] Weiwen Dai,[‡] Dieter Brömme,[||] D. Margaret Worrall,[⊥] Gary A. Silverman,[@] Theresa H. T. Coetzer,[#] Clive Dennison,[#] Stephen P. Bottomley,[‡] and James C. Whisstock*[,‡,+]

Department of Biochemistry and Molecular Biology, Monash University Clayton Campus, P.O. Box 13D, Victoria 3800, Australia, Department of Human Genetics, Mt. Sinai School of Medicine, New York, New York 10029, Department of Biochemistry and Conway Institute, University College Dublin, Belfield, Dublin 4, Ireland, Department of Pediatrics, Harvard Medical School, Children's Hospital, 300 Longwood Avenue, Enders 970, Boston, Massachusetts 02115, School of Molecular and Cellular Biosciences-Biochemistry, University of Natal, Private Bag X01, Scottsville 3209, South Africa, and Victorian Bioinformatics Consortium, Monash University, Victoria 3800, Australia

Received December 3, 2001; Revised Manuscript Received February 4, 2002

ABSTRACT: The closely related serpins squamous cell carcinoma antigen-1 and -2 (SCCA-1 and -2, respectively) are capable of inhibiting cysteine proteases of the papain superfamily. To ascertain whether the ability to inhibit cysteine proteases is an intrinsic property of serpins in general, the reactive center loop (RCL) of the archetypal serine protease inhibitor α_1-antitrypsin was replaced with that of SCCA-1. It was found that this simple substitution could convert α_1-antitrypsin into a cysteine protease inhibitor, albeit an inefficient one. The RCL of SCCA-1 is three residues longer than that of α_1-antitrypsin, and therefore, the effect of loop length on the cysteine protease inhibitory activity was investigated. Mutants in which the RCL was shortened by one, two, or three residues were effective inhibitors with second-order rate constants of 10^5-10^7 M^{-1} s^{-1}. In addition to loop length, the identity of the cysteine protease was of considerable importance, since the chimeric molecules inhibited cathepsins L, V, and K efficiently, but not papain or cathepsin B. By testing complexes between an RCL-mimicking peptide and the mutants, it was found that the formation of a stable serpin−cysteine protease complex and the inhibition of a cysteine protease were both critically dependent on RCL insertion. The results strongly indicate that the serpin body is intrinsically capable of supporting cysteine protease inhibition, and that the complex with a papain-like cysteine protease would be expected to be analogous to that seen with serine proteases.

Serpins are single-domain proteins found in plants, animals, and viruses, which participate in a wide variety of intracellular and extracellular processes (*2*, *3*). Most characterized serpins display inhibitory activity against trypsin-like or chymotrypsin-like serine proteases (*4*), but the family includes some noninhibitory members. In addition, inhibitory activity against "nonconventional" targets has been demonstrated in certain cases. These include caspases (*5*), subtilisins (*6*), the aspartic protease pepsin (*7*), and, most recently, cysteine proteases of the papain superfamily (*8*, *9*). Serpins are comprised of three β-sheets, eight or nine α-helices, and a reactive center loop (RCL)[1] through which the initial interaction with a cognate protease occurs. The RCL amino acid residues are presented as a pseudosubstrate, with cleavage initiated between the P_1 and P_1' residues (refer to footnote 1 for an explanation of P_n notation). During the reaction with a serine protease, the serpin undergoes a profound conformational change, culminating in the translocation of the protease to the opposite end of the molecule (~70 Å) (*10*), yielding a protease−serpin complex irreversibly and covalently trapped via an acyl ester bond (*11*, *12*). The precise mechanism of cysteine protease inhibition remains to be characterized.

Squamous cell carcinoma antigen-1 (SCCA-1), a member of the intracellular serpin branch of the superfamily (*13*, *14*), is an inhibitor of the cysteine proteases cathepsins K, L, and

† This work was supported by grants from the Australian Research Council (A00103027 to R.N.P. and J.C.W. and A10017123 to S.P.B. and R.N.P.). J.C.W. and S.P.B. are Monash University Logan Fellows, and J.C.W. is a National Health and Medical Research Council Senior Research Fellow.

* To whom correspondence should be addressed: Department of Biochemistry and Molecular Biology/VBC, Monash University Clayton Campus, P.O. Box 13D, Victoria 3800, Australia. Telephone: +61 3 9905 3747. Fax: +61 3 9905 4699. E-mail: James.Whisstock@med.monash.edu.au.

‡ Department of Biochemistry and Molecular Biology, Monash University.

§ These authors contributed equally to this work.

|| Mount Sinai School of Medicine.

⊥ University College Dublin.

@ Children's Hospital.

University of Natal.

+ Victorian Bioinformatics Consortium, Monash University.

[1] Abbreviations: AT, α_1-antitrypsin, or SERPINA1; cysteine protease, cysteine proteases of the papain (C1) superfamily; RCL, reactive center loop; SAT, SAT-1, SAT-2, etc., chimeric mutants of α_1-antitrypsin in which the reactive center loop has been replaced with that from SCCA-1; SCCA-1, squamous cell carcinoma antigen-1, or SERPINB3; SCCA-2, squamous cell carcinoma antigen-2, or SERPINB4. According to the convention of Schechter and Berger (*1*), the residues of a peptide substrate are designated P_n-P_2, P_1, and $P_1'-P_n'$ and interact with corresponding subsites in the protease designated S_n-S_2, S_1, and $S_1'-S_n'$, respectively; cleavage occurs by definition between the P_1 and P_1' positions.

10.1021/bi0159985 CCC: $22.00
Published on Web 03/19/2002

S and papain (*8*). The sequence of SCCA-1 is 92% identical with that of SCCA-2, which has poor anti-cysteine protease activity. This activity was, however, increased 50-fold upon exchange of the SCCA-2 RCL for that of SCCA-1 (*15*). While this study confirmed the importance of the RCL amino acid sequence to cysteine protease inhibition, it did not preclude the possibility that extrinsic amino acids play a role: the two serpins have arisen from a recent gene duplication (*13*), and thus, their activity may represent a unique evolutionary adaptation of serpin function. Only one other serpin−papain-like cysteine protease interaction has been characterized, between the more distantly related serpin, antithrombin, and papain. Importantly, inhibition of the latter occurred in a distinctly "non-serpin-like" fashion, with a poor second-order association rate constant and a high stoichiometry of inhibition, and involving degradation of the "trapped" protease by active, noncomplexed papain molecules (*9*). It is therefore unclear, due in part to striking differences in protease catalytic machinery and architecture, whether inhibitors of serine and cysteine proteases have special structural or sequence requirements for engaging in effective (rapid) and efficient (close to equimolar) inhibition of a papain-like cysteine protease molecule.

We set out to determine whether the serpin scaffold is intrinsically capable of supporting such effective and efficient cysteine protease inhibition per se, and selected the well-characterized and phylogenetically distinct serpin α_1-antitrypsin (AT) for this purpose. AT is the physiological antagonist of the serine protease elastase (*16*). It arises from a different evolutionary branch of the serpin superfamily with respect to SCCA-1 (*3*), and most of the amino acids common to both AT and SCCA-1 correspond with residues generally conserved throughout the serpin superfamily. Thus, observations made with AT about the requirements for cysteine protease inhibition should be applicable to the cysteine protease-inhibiting potential of serpins generally.

We show here that despite a low level of primary sequence identity between SCCA-1 and AT (33%), chimeras of the latter are capable of supporting highly efficient cysteine protease inhibition, in a manner influenced by the RCL length. By examining interactions between proteases and complexes of the mutants with an RCL-mimicking peptide, we show that, as with serine proteases, inhibition requires insertion of the RCL into the body of the molecule.

EXPERIMENTAL PROCEDURES

Materials. The pQE-31 vector containing the sequence of the mature region of human α_1-antitrypsin, pQE-31/AT, was a generous gift from L. Jin (*17*). DNA-modifying enzymes were from Boehringer Mannheim (Mannheim, Germany). Sheep cathepsin L (*18*) and human cathepsins K and V (*19*, *20*) were prepared as described previously. Bovine cathepsin B and all fine chemicals were purchased from Sigma (Sydney, Australia), and papain was purchased from ICN (Sydney, Australia). Oligonucleotides were from Genset Pacific (Lismore, Australia). Sepharose resins and HiTrap columns were from Pharmacia (Uppsala, Sweden).

Plasmid Construction. The pQE-31/AT vector, containing *Pst*I−*Ava*I restriction sites flanking the P_{12}−P_3' codons of the RCL, was modified to render these sites unique through digestion and religation of a single *Xho*I site and *Sph*I−*Hin*dIII sites immediately downstream of the AT insert. Oligonucleotide pairs were synthesized encoding portions of the RCL of SCCA-1 (as shown in Figure 1), annealed by slow cooling from 95 °C, phosphorylated using T4 polynucleotide kinase, and ligated into *Pst*I/*Ava*I-cut pQE31-AT. Successful mutagenesis was confirmed using diagnostic digests and automated sequencing, and plasmid DNA was used to transform JM101 cells for expression. The resulting chimeras were designated SAT, SAT-1, SAT-2, and SAT-3 (Figure 1).

```
AT          EKGTEAAGAMFLEAIPMSIP---PE
SCCA-1      EEGAEAAAAATAVVGFGSSPASTNEE
SAT         EKGTEAAAATAVVGFGSSPASTNPE
SAT-1       EKGTEAAAATAVVGFGSSPAST-PE
SAT-2       EKGTEAAAATAVVGFGSSPAS--PE
SAT-3       EKGTEAAAATAVVGFGSSPA---PE
```

FIGURE 1: RCL sequences of wild-type AT (AT), SCCA-1 (SCCA-1), and the four mutants used in this study (SAT−SAT-3). P_1−P_1' residue pairs are highlighted in gray; residue differences with respect to SCCA-1 are highlighted in black.

Protein Expression and Purification. To express AT and mutants, 1 L of 2YT medium (containing 10 μg/mL ampicillin) was inoculated with an overnight culture and grown to an OD_{600} of 0.6−0.8 prior to induction with 0.5 mM IPTG. The cultures were incubated with shaking at 37 °C for 5 h, and the harvested cell pellet was resuspended in buffer A [25 mM $NaPO_4$ and 0.5 M NaCl (pH 8.0)] and disrupted by sonication. The soluble fraction was applied to a 5 mL nickel-charged NTA-Sepharose resin and washed to baseline in buffer A with 25 mM imidazole. The bound protein was eluted using buffer A with 250 mM imidazole, then diluted 10-fold into buffer B [50 mM Tris (pH 8.0) and 10 mM NaCl], and applied to a 20 mL Q-Sepharose column. After being washed with buffer B to baseline, the protein was eluted using a 10 to 400 mM NaCl gradient in a volume of 100 mL. Fractions of the intact, monomeric protein were selected on the basis of Coomassie-stained denaturing SDS−PAGE (*21*) and native PAGE gels (*22*). Where necessary, fractions were pooled and bound to a 5 mL glutathione−Sepharose column [previously charged using 0.1% (w/v) 5,5′-dithiobis(2-nitrobenzoic acid), 0.1 M Tris, and 5 mM EDTA (pH 8.0)], washed with 0.1 M Tris, 0.5 M NaCl, and 5 mM EDTA (pH 8.0), and eluted with 5 mM 5,5′-dithiobis(2-nitrobenzoic acid), 5 mM dithiothreitol, 0.1 M Tris, and 5 mM EDTA (pH 8.0). SCCA-1, in the pRSET vector system (Invitrogen, CA), was expressed in BL21 cells in the same fashion and, after elution from the NTA-Sepharose resin, was diluted 5-fold into 50 mM MES and 50 mM NaCl (pH 5.9) and loaded onto a 5 mL HiTrap SP column. A 60 mL salt gradient was applied into 50 mM MES and 1 M NaCl (pH 5.9), and fractions were selected following SDS−PAGE and native PAGE analysis.

Enzyme Assays. Inhibitor assays were performed in cathepsin buffer [0.1 M acetate (pH 5.5), 1 mM EDTA, 0.1% (w/v) Brij35, and 10 mM cysteine] at 30 °C on a Fluorostar fluorimetric plate reader (BMG Labtechnologies), using *N*-cbz-Phe-Arg-methylcoumarin substrate (Sigma). To determine the stoichiometry of inhibition (SI), various concentrations of inhibitor were incubated with a constant amount of protease at 30 °C for 45 min, after which the substrate was added to a concentration of 10 μM and residual

protease activity measured. The rate of substrate turnover (change in fluorescence per unit time) was plotted against the molar ratio of inhibitor to enzyme and a linear regression analysis performed. The intercept with the abscissa yielded the SI value. The second-order association rate constant (k_{ass}) was determined using a continuous assay under pseudo-first-order conditions (*23*), in the presence of 1% (w/v) PEG 8000 and 50 or 100 μM substrate in cathepsin buffer. Within an assay, enzyme concentrations were held constant at a value between 0.01 and 1 nM, variable inhibitor concentrations between 10- and 1000-fold greater than that of the enzymes were used, and the total volume was 200 μL. Inhibitor-free controls indicated the enzymes remained at least 95% active under the conditions that were used, and the extent of substrate depletion was less than 5%. Upon addition of enzyme to an inhibitor/substrate mixture, cleaved fluorogenic substrate fluorescence was monitored continuously using excitation and emission wavelengths of 370 and 460 nm, respectively. These fluorescence data were fitted, using least-squares regression in PRISM (GraphPad Software), to an equation describing slow, tight inhibitor binding (*24*):

$$F = V_s t + (V_0 - V_s)\frac{1 - e^{-k_{obs}t}}{k_{obs}}$$

where F is the absolute fluorescence value at time t, V_0 is the velocity at time zero, V_s is the residual velocity once the reaction has run to completion, and k_{obs} is the apparent rate constant. The uncorrected second-order association rate constant (k_{unc}) was obtained from the slope of the linear regression through a plot of k_{obs} versus inhibitor concentration. The substrate-corrected rate constant was obtained using the equation

$$k_{ass} = k_{unc}\left(1 + \frac{[S]}{K_m}\right)$$

The final, substrate-corrected k_{ass} values were calculated using K_m values determined under equivalent assay conditions: 9.8 μM ovine cathepsin L (*25*), 59 μM human cathepsin K, 27 μM human cathepsin V, and 160 μM papain (*26*).

Identification of the Scissile Bond. SAT-2 (2 μM) was reacted with 0.56 μM cathepsin L for 30 min at 25 °C in a 50 μL volume of cathepsin buffer, run on a 15% SDS−PAGE Tris-Tricine gel, and transferred onto a PVDF membrane. The band corresponding to the released 7 kDa fragment was excised and submitted for N-terminal sequencing (Department of Biochemistry, LaTrobe University, Bundoora, Australia).

Formation of the Binary Complex. SAT-2 or SAT-3 (10 μM), in buffer B, was incubated in the presence of 1 mM 12mer peptide (Auspep), corresponding to the $P_{14}-P_3$ reactive center loop sequence of antithrombin, at 4 °C for 72 h. Complete formation of the binary complex between the chimeras and peptide was confirmed using native PAGE (*27*). Control and peptide-complexed material was then assayed for activity against cathepsin V at a 100-fold excess, under the conditions described above.

SDS−PAGE Gels. SAT and variants, at a final concentration of 10 μM in 20 μL of cathepsin buffer, were incubated with equimolar amounts of sheep cathepsin L or bovine cathepsin B for 10 min at 25 °C, prior to addition of 10 μM E-64, boiling in Laemmli loading buffer (*21*), and separation on a 12.5% SDS−PAGE gel.

Native Acid PAGE Gels. SAT-2 (5.3 μM), or antithrombin peptide-complexed SAT-2, was incubated with 3.1 μM cathepsin V or 0.6 μM cathepsin B, in a final volume of 20 μL of cathepsin buffer at 25 °C for 10 min. Samples, with appropriate controls, were electrophoresed on a native PAGE gel under mildly acidic conditions, at 4 °C and a constant voltage of 80 V. A discontinuous gel system based on that used by Zhou et al. (*28*) was developed, with 40 mM β-alanine/40 mM imidazole anode running buffer and 40 mM bis-tris cathode running buffer, adjusted to pH 5.8 with acetic acid. The stacking gel contained 2.5% (w/v) acrylamide, 0.08% (w/v) bisacrylamide, and 90 mM acetic acid, adjusted to pH 6.0 with KOH, and the running gel contained 10% (w/v) acrylamide, 0.33% (w/v) bisacrylamide, and 90 mM acetic acid, adjusted to pH 5.8 with KOH. Methyl green was used as the tracking dye, and the protein migrated from the anode to the cathode.

Table 1: Inhibition of Papain-like Cysteine Proteases by Different AT Mutants[a]

	SAT	SAT-1	SAT-2	SAT-3	SCCA-1
ovine cathepsin L					
SI	57.2	6.7	3.6	4.8	1.0
k_{ass}[b]	−	4.96×10^5	9.84×10^5	1.32×10^6	1.37×10^6
human cathepsin V					
SI	19.4	1.6	1.7	1.7	1.0
k_{ass}[b]	−	1.19×10^5	2.35×10^5	3.20×10^5	2.47×10^6
human cathepsin K					
SI	47.5	1.2	2.3	6.4	∼1[c]
k_{ass}[b]	−	9.29×10^5	1.38×10^6	9.89×10^5	8.76×10^6
bovine cathepsin B					
SI	ni	≪	≪	≪	ni[c]
papain					
SI	ni	ni	ni	ni	1.3
k_{ass}[b]	−	−	−	−	1.78×10^5

[a] The stoichiometry of inhibition (SI) and, where determined, the second-order rate constant (k_{ass}) are shown. All standard errors were less than 5%. ni means no inhibition detected; ≪ means little inhibitory activity observed. [b] k_{ass} in units of M^{-1} s^{-1}. [c] SI determined by Schick et al. (*8*).

RESULTS

To determine whether a serpin evolutionarily removed from SCCA-1 and -2 could support irreversible inhibition of papain-like cysteine proteases, we constructed RCL mutants of AT. Replacement of 12 amino acids in AT (P_9-P_3') with 15 amino acids from the RCL of SCCA-1 (P_9-P_6') yielded the first chimera, SAT, and further mutants were generated by progressively shortening the C-terminal (or P′) portion of the introduced loop: P_9-P_5', SAT-1; P_9-P_4', SAT-2; and P_9-P_3', SAT-3 (Figure 1). Wild-type AT failed to inhibit any of the proteases considered in this study.

SAT Inhibitory Activity. The AT mutant bearing the full-length RCL of SCCA-1 (SAT) was a poor inhibitor of cathepsins L, V, and K, with a stoichiometry of inhibition (SI) 19−57-fold greater than the value of 1 seen with SCCA-1 (Table 1, columns 3 and 7). The SI value reflects the balance between a serpin molecule adopting substrate or inhibitor behavior at the key "decision point" in the reaction pathway. A high SI is an indication of an inhibitor more prone to substrate-like behavior, as with the SAT

interactions. SAT failed to inhibit either cathepsin B or papain.

A significant improvement in SI was seen with the removal of the P_6' residue, yielding an introduced RCL two amino acids longer than that of AT (SAT-1). SAT-1 was an efficient inhibitor of cathepsin K, with almost all molecules undergoing a productive, inhibitory interaction. This was followed closely by cathepsin V, while the interaction with cathepsin L was still relatively poor (Table 1, column 4); SAT-1 failed to significantly inhibit either cathepsin B or papain.

The second-order association rate constant (k_{ass}) reflects the rate of the inhibitory interaction and, hence, is a measure of the effectiveness of an inhibitor. Progress curves, obtained by continually monitoring the loss of enzyme activity, were found to fit the standard equation describing a slow, tight binding interaction dependent (under pseudo-first-order conditions) only on inhibitor concentration (*24*). An inhibitory pathway requiring digestion of trapped protease would be expected to show a further dependence on enzyme concentration, as shown by Bjork et al. (*9*). However, when the protease concentration was varied 10-fold, this did not affect the observed rate constant. From this equation, k_{ass} values for the three inhibited enzymes and SAT-1 were calculated, ranging from 1×10^5 to 9×10^5 M^{-1} s^{-1}. These values are comparable to that seen with some physiological interactions [e.g., inhibition of granzyme B by CrmA (*29*)], although 2.5−9-fold lower than seen with SCCA-1 and the respective proteases (Table 1, columns 4 and 7).

Removal of a further P_5' residue yielded SAT-2, which had varied effects on the SI and k_{ass} of the inhibitory interaction. A shorter RCL led to an improved SI against cathepsin L but a poorer SI with cathepsin K and did not alter that seen with cathepsin V (Table 1, column 5). Interestingly, while there was only a marginal increase in k_{ass} against cathepsins V and K when compared with that of SAT-1, the k_{ass} seen with this mutant for cathepsin L doubled to give a value very similar to that of the wild-type SCCA-1−cathepsin L pair.

Deletion of the P_4' residue in SAT-3, whose RCL is the same length as that of AT, had a deleterious effect on the SI of the interaction with cathepsins L and K, but did not affect the SI with cathepsin V, nor did it lead to markedly different k_{ass} values (Table 1, column 6).

SCCA-1 Inhibitory Activity. Second-order association rate constants for the interaction of SCCA-1 and cathepsins K and L of 1.1×10^5 and 3.0×10^5 M^{-1} s^{-1} have been determined previously at 25 °C under slightly different buffer conditions (*8*). Under the conditions used here, rate constants were found to be greater by almost 10-fold than those reported, and SCCA-1 was found to be a more effective inhibitor of cathepsin K than of cathepsin L (Table 1, column 7). The data presented here show that SCCA-1 also inhibits cathepsin V with inhibitory parameters similar to that seen with cathepsin L, while the inhibitory interaction with papain was slower and less efficient than with the other three enzymes.

Scissile Bond. The C-terminal fragment generated by the interaction of cathepsin L and SAT-2 had the sequence SSPASPE, confirming that SAT-2 was cleaved at the same site in its RCL as SCCA-1 (*8*).

SDS−PAGE Gels. SDS−PAGE analysis of the interaction between the SAT mutants and cathepsin L showed that

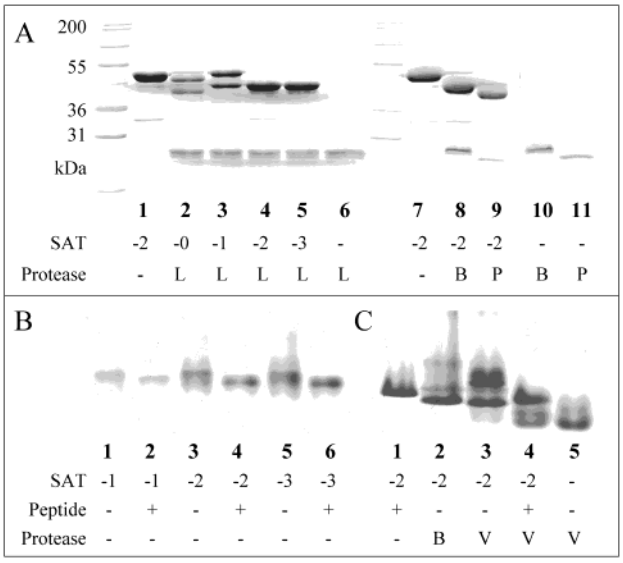

FIGURE 2: (A) SDS−PAGE gel illustrating the interaction between SAT variants and cathepsin proteases. Molecular mass markers are shown at the left. The mutants, designated -0 for SAT, -1 for SAT-1, etc., were incubated for 10 min with a 1:1 ratio of sheep cathepsin L (L), bovine cathepsin B (B), or papain (P). SAT-2 alone and protease alone (L, B, or P) was used in reference lanes. (B) Native PAGE gel of the binary complex. The SAT variants (SAT-1, -2, and -3) were incubated for 72 h at 4 °C with (+) or without (−) a peptide mimicking the $P_{14}-P_3$ residues of antithrombin. A faster-migrating (lower) band indicates incorporation of the acetylated peptide. (C) Native PAGE gel of the SAT-2 complex under acidic conditions. SAT-2, in two cases preincubated with the antithrombin RCL peptide (+), was incubated with a 1.7:1 ratio of cathepsin V (V) or a 1:0.1 ratio of cathepsin B (B) for 10 min at room temperature.

SAT-2 and SAT-3 were fully susceptible to RCL cleavage during the inhibitory process (Figure 2A, lanes 4 and 5). However, SAT and SAT-1 behaved in a nonideal fashion. Alternative cleavage products of the serpin were seen in the interaction between SAT and cathepsin L, which may reflect inappropriate RCL presentation to the protease (lane 2), and suggests an explanation for its predominantly substrate-like behavior. SAT-1 cleavage by cathepsin L generated a product of the expected size, although approximately two-thirds of the SAT-1 molecules remained uncleaved (lane 3). The RCL of SAT-1 would be expected to be fully available given that it interacts in a ratio of almost 1:1 with cathepsin K (Table 1, column 4); it may be that the RCL exists in two or more alternative conformations and that while the active site topology of cathepsin K accommodates these different forms, cathepsin L does not.

Despite the lack of inhibition of cathepsin B by the SAT mutants, SAT-2 was found to be cleaved in the appropriate place (Figure 2A, lane 8). This demonstrates that a lack of inhibitory activity is not due to the inability of cathepsin B to accommodate the RCL and form an initial noncovalent docking complex, but must reflect dysfunction later in the inhibitory pathway. This also appears to be the case with papain. While papain apparently cleaved the RCL of SAT-2 in a different position (Figure 2A, lane 9), a time course assay revealed that this was a secondary cleavage product and that the RCL was initially cleaved in the appropriate position (gel not shown).

It is noteworthy that the band corresponding to cathepsin L (lane 6) showed a complete lack of protease degradation

with any of the four mutants (Figure 2A, lanes 2−5), which distinguishes the inhibitory interaction from that seen between papain and antithrombin, in which the protease was found to be extensively digested (*9*).

Properties of the Binary Complex. Using native PAGE analysis, SAT-1, SAT-2, and SAT-3 (Figure 2B, lanes 1, 3, and 5) were found to be fully capable of forming a binary complex with a 12mer peptide mimicking the $P_{14}-P_3$ RCL residues of antithrombin (lanes 2, 4, and 6). This demonstrates that these preparations were not contaminated with latent, inactive material, since latent AT is not able to incorporate an exogenous peptide (*30*).

Formation of a binary complex between SAT-2 or SAT-3 and the 12mer antithrombin peptide was found to completely abolish inhibitory activity against cathepsin V. Crystallographic and other data have shown that an RCL-mimicking peptide is inserted into the center of the A β-sheet (*31*), preventing insertion of the RCL and, in turn, inhibition of serine proteases. The results presented here demonstrate that inhibitory activity against a cysteine protease is dependent on the ability of the A β-sheet to accommodate the protease-bound RCL.

Acid Native PAGE Gel. To confirm that a stable, rather than transient, interaction takes place between a serpin and a papain-like cysteine protease, the complex between SAT-2 and cathepsin V was electrophoresed in an acid native PAGE system (Figure 2C). Incubation of SAT-2 and cathepsin V at a ratio of 1.7:1, the stoichiometry of the interaction, yielded a slower-migrating (higher) band corresponding to the inhibitor−enzyme complex, a lower band corresponding to nonproductively cleaved material, and a disappearance of the cathepsin V band (lane 3). In contrast, peptide-complexed SAT-2 (lane 1) failed to form a slower-migrating complex when incubated with cathepsin V under the same conditions (lane 4), indicating that the formation of a stable complex is dependent on RCL insertion.

DISCUSSION

In this study, we constructed a number of variants of AT bearing the full-length, and shortened, portions of the RCL of SCCA-1. The data show that AT is, given the appropriate RCL sequence, entirely capable of efficient cysteine protease inhibition.

While the long RCL of SCCA-1 appears to perturb the inhibitory mechanism in SAT, reflected by a high stoichiometry of inhibition (SI), shorter lengths (SAT-1, -2, and -3) lead to a marked improvement in the efficiency of the interaction. The data presented here demonstrate that the SI of the SAT−cysteine protease reaction is altered by (1) the identity of the target enzyme and (2) the length of the transplanted RCL. The RCL length did not have a marked influence on the second-order association rate constant (k_{ass}), which was found to be similar, for a given protease, among the three inhibitors. This similarity, despite the differing SI values, suggests that differences in loop length have a minimal effect on the kinetics of the interaction between the RCL and protease, but strongly affect the balance between the inhibitor and substrate pathways in the inhibitory reaction scheme.

The data also show, however, that a simple exchange of RCL will not make AT into SCCA-1: only the pairing of cathepsin K and SAT-1 yielded a SI value comparable to that seen with SCCA-1, and only SAT-2 or SAT-3 interacting with cathepsin L yielded a comparable k_{ass}. It has been observed with serpin−serine protease pairs that the effectiveness of a chimera depends on both the RCL sequence and the serpin scaffold (*32*, *33*). In addition, variation in the length of the RCL either N-terminal to (*17*) or C-terminal to (*32*) the scissile bond has been shown to produce varied—though in most cases deleterious—results, depending somewhat on the identity of the target protease. Given that there is no mutant that achieves a 1:1 interaction for all enzymes, RCL length in the C-terminal (P′) residues is not by itself able to account for the generally increased SI values.

In most cases, the position of the P_1-P_1' bond, where the serpin is cleaved by its target protease, is 17 amino acids downstream from the start of the highly conserved hinge region motif ($EEGTEAAAX_8P_1-P_1'$). In the serpin−enzyme complex, the RCL is fully inserted into the A β-sheet of the serpin, with the P_1 residue positioned at the base of the molecule, covalently linked by an acyl bond to the active site serine residue of the protease. It has been shown that the compression of the enzyme against the serpin distorts the active site (*10*, *34*). We have shown here, using peptide-complexed serpin, that inhibition of cysteine proteases, and formation of a stable serpin−enzyme complex, requires RCL insertion. Additionally, papain has been demonstrated by others to become susceptible to proteolysis following interaction with antithrombin, suggestive of distortion of the molecule (*9*). Taken together, this evidence suggests that the final complex will be similar in nature to that seen with serine proteases. It is therefore noteworthy that the P_1-P_1' residue in SCCA-1 is 16 residues from the start of the hinge motif (*8*), making the N-terminal portion of the RCL one residue shorter than that of AT and most other serpins (Figure 1). The importance of this N-terminal length is highlighted, in the case of crmA, by rearrangements in the loops near the base of the A sheet, which accommodate an inserted RCL of the same length as that of SCCA-1 (*35*). It may be that the base of SCCA-1 has also evolved a local structure more suited to a shorter inserted N-terminal loop length. Without such adjustments, tethering of the protease to a shortened loop might lead to an increase in the number of steric clashes, which may cause the protease active site topology to exert a stronger influence over the mechanics of the inhibitory reaction, and hence lead to the variation in SI, from protease to protease, seen with the AT mutants.

It is noteworthy that, while the SI of the interaction between SAT-1 and cathepsin L was high (6.7), only a proportion of the serpin was found to be cleaved (Figure 2A). In contrast, almost all of the SAT-1 molecules interacted with cathepsin K, giving an SI value close to 1. This suggests an interesting scenario in which loop presentation in a subset of the SAT-1 molecules is incompatible with the active site of cathepsin L, but still permits docking into the active site of cathepsin K (Table 1). A previous study has shown that the length of the N-terminal (P) side of the RCL is capable of substantially influencing the SI (*17*). In contrast, another study in which the C-terminal (P′) side of the RCL of antitrypsin was lengthened showed a much smaller variation in SI values (*32*). This is not surprising, given that it is the P side that is covalently tethered to the protease and is inserted into the central A sheet, and thus is integral to the

serpin mechanism. In this study, we show that the length of the P′ region is capable of dramatically influencing the SI of the interaction between serpins and certain cysteine proteases. While it is expected that the P′ side plays no part in the inhibitory mechanism following RCL cleavage, P′ residues may exert an influence over loop conformation and hence the presentation of the region surrounding the scissile bond. This appears to be the case with SAT-1, and suggests one way in which serpins may discriminate between two enzymes having ostensibly similar subsite specificities. The potential impact of this observation is particularly evident in light of the marked variation seen in the length of the P′ region when the serpin superfamily is viewed as a whole (*3*).

The interaction with cathepsin B and papain is also interesting in this regard. While both were able to recognize and cleave the RCL in an appropriate fashion, no inhibitory complex was formed. Cathepsin B is slightly different from the other cysteine proteases considered in this study. The presence of an extra, "occluding" loop in the P′ region of the active site may have some role to play, perhaps by stabilizing the active site geometry against deformation, or by presenting a steric barrier which favors dissociation of the complex during, or following, loop insertion. It is clear that the occluding loop does not impede initial docking of the RCL with cathepsin B, since the SAT-2−cathepsin B interaction leads to the appearance of RCL-cleaved serpin (Figure 2A). Another difference evident from the crystal structure of cathepsin B [PDB entry 1huc (*36*)] is the presence of additional disulfide bonds, which may stabilize the structure against deformation-mediated inhibition. In particular, the bridge between Cys14 and Cys43 (cathepsin B numbering), which is not present in cathepsins K, L, and V or papain, tethers the two lobes of the enzyme together at the base of the A helix, on the side of the molecule opposite the active site cleft. The active site cysteine is situated on one lobe of the enzyme and the histidine−asparagine pair on the other. Any inhibitory interaction dependent on a gross distortion of the molecule via movement of the two lobes would therefore be less likely to occur.

Papain is more closely related to cathepsins K, L, and V than to cathepsin B, and is inhibited by SCCA-1 [Table 1 (*8*)]; it is therefore surprising that no inhibitory activity was observed. Initially, SDS−PAGE analysis of the SAT-2 cleavage product suggested that inappropriate proteolysis, at a site N-terminal to that expected, had occurred. A time course experiment revealed that this was in fact a secondary cleavage product, and that the (noninhibitory) interaction initially produced a cleaved serpin of the expected size. No features were observed in the structure of papain [PDB entry 1pap (*37*)] which would account for its resistance to inhibition, although it may be a more stable enzyme, with 9% more hydrogen bonds and 200% more salt bridges than human cathepsin L.

Two mechanisms for the inhibition of cysteine proteases of the papain superfamily by serpins have been proposed. In considering the interaction between papain and antithrombin, Björk et al. (*9*) suggested that distortion of trapped papain permitted proteolytic degradation by free protease. In this study, SDS−PAGE gels failed to show any degradation of the proteases (Figure 2A), suggesting that inhibition occurred in the more conventional manner seen with SCCA-1 (*8*). While we failed to see a complex on a Coomassie-stained SDS−PAGE gel, it is worthwhile noting that in the latter study a Western blot was required for this purpose. Additionally, no complex has been seen with the interaction between caspases and crmA using SDS−PAGE gels (*38*), and it should also be noted that a thioester bond, the covalent bond that would be expected to link cysteine protease and serpin, is more susceptible to nucleophilic attack than the ester bond seen with serine proteases (*39*). We were, however, able to visulize the serpin−cysteine protease complex under nondenaturing conditions (Figure 2C).

We have shown here that the body of a typical serine protease-inhibiting serpin is capable of supporting effective cysteine protease inhibition, upon substitution of the RCL of a cysteine protease inhibitor from a different clade of the serpin superfamily. This suggests the serpin body is not an impediment to the inhibition of papain-like cysteine proteases; therefore, such an ability could be expected to arise where the opportunity and necessity exist. The data suggest that there are peripheral adjustments in SCCA-1 which have "fine-tuned" its interaction with cysteine proteases. In particular, there may be accommodating rearrangements in the base of SCCA-1, by analogy with crmA, given its shorter N-terminal RCL length. It is clear, however, that structural differences between proteases also exert an effect; we have identified several structural features in cathepsin B which may explain its resistance to inhibition by SCCA-1 and the antitrypsin mutants. The physiological role of papain-like cysteine protease inhibition by serpins is not yet known. If cathepsin B is shown to be generally resistant to serpin-mediated inhibition, this would begin to indicate the boundaries of that role, in relation to other inhibitors such as cystatins.

REFERENCES

1. Schecter, I., and Berger, A. (1967) *Biochem. Biophys. Res. Commun. 27*, 157−162.
2. Silverman, G. A., Bird, P. I., Carrell, R. W., Church, F. C., Coughlin, P. B., Gettins, P. G., Irving, J. A., Lomas, D. A., Luke, C. J., Moyer, R. W., Pemberton, P. A., Remold-O'Donnell, E., Salvesen, G. S., Travis, J., and Whisstock, J. C. (2001) *J. Biol. Chem. 276*, 33293−33296.
3. Irving, J. A., Pike, R. N., Lesk, A. M., and Whisstock, J. C. (2000) *Genome Res. 10*, 1845−1864.
4. Gettins, P. G. (2000) *Genome Res. 10*, 1833−1835.
5. Komiyama, T., Ray, C. A., Pickup, D. J., Howard, A. D., Thornberry, N. A., Peterson, E. P., and Salvesen, G. (1994) *J. Biol. Chem. 269*, 19331−19337.
6. Komiyama, T., Gron, H., Pemberton, P. A., and Salvesen, G. S. (1996) *Protein Sci. 5*, 874−882.
7. Mathialagan, N., and Hansen, T. R. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 13653−13658.
8. Schick, C., Pemberton, P. A., Shi, G. P., Kamachi, Y., Cataltepe, S., Bartuski, A. J., Gornstein, E. R., Bromme, D., Chapman, H. A., and Silverman, G. A. (1998) *Biochemistry 37*, 5258−5266.
9. Bjork, I., Nordling, K., Raub-Segall, E., Hellman, U., and Olson, S. T. (1998) *Biochem. J. 335*, 701−709.
10. Huntington, J. A., Read, R. J., and Carrell, R. W. (2000) *Nature 407*, 923−926.
11. Lawrence, D. A., Ginsburg, D., Day, D. E., Berkenpas, M. B., Verhamme, I. M., Kvassman, J. O., and Shore, J. D. (1995) *J. Biol. Chem. 270*, 25309−25312.
12. Egelund, R., Rodenburg, K. W., Andreasen, P. A., Rasmussen, M. S., Guldberg, R. E., and Petersen, T. E. (1998) *Biochemistry 37*, 6375−6379.

13. Schneider, S. S., Schick, C., Fish, K. E., Miller, E., Pena, J. C., Treter, S. D., Hui, S. M., and Silverman, G. A. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 3147−3151.
14. Barnes, R. C., and Worrall, D. M. (1995) *FEBS Lett. 373*, 61−65.
15. Luke, C., Schick, C., Tsu, C., Whisstock, J. C., Irving, J. A., Bromme, D., Juliano, L., Shi, A. U. G. P., Chapman, H. A., and Silverman, G. A. (2000) *Biochemistry 39*, 7081−7091.
16. Gadek, J. E., Fells, G. A., Zimmerman, R. L., Rennard, S. I., and Crystal, R. G. (1981) *J. Clin. Invest. 68*, 889−898.
17. Zhou, A., Carrell, R. W., and Huntington, J. A. (2001) *J. Biol. Chem. 276*, 27541−27547.
18. Dennison, C., Moolman, L., Pillay, C. S., and Meinesz, R. E. (2000) *Anal. Biochem. 284*, 157−159.
19. Bromme, D., Li, Z., Barnes, M., and Mehler, E. (1999) *Biochemistry 38*, 2377−2385.
20. Linnevers, C. J., McGrath, M. E., Armstrong, R., Mistry, F. R., Barnes, M. G., Klaus, J. L., Palmer, J. T., Katz, B. A., and Bromme, D. (1997) *Protein Sci. 6*, 919−921.
21. Laemmli, U. K. (1970) *Nature 227*, 680−685.
22. Bryan, J. K. (1977) *Anal. Biochem. 78*, 513−519.
23. Le Bonniec, B. F., Guinto, E. R., and Stone, S. R. (1995) *Biochemistry 34*, 12241−12248.
24. Rovelli, G., Stone, S. R., Guidolin, A., Sommer, J., and Monard, D. (1992) *Biochemistry 31*, 3542−3549.
25. Coetzer, T. H., Dennehy, K. M., Pike, R. N., and Dennison, C. (1995) *Comp. Biochem. Physiol., Part B: Biochem. Mol. Biol. 112*, 429−439.
26. Irving, J. A., Shushanov, S. S., Pike, R. N., Popova, E. Y., Bromme, D., Coetzer, T. H., Bottomley, S. P., Boulynko, I. A., Grigoryev, S. A., and Whisstock, J. C. (2002) *J. Biol. Chem.* (in press).
27. Chang, W. S. W., Whisstock, J. C., Hopkins, P. C. R., Lesk, A. M., Carrell, R. W., and Wardell, M. R. (1997) *Protein Sci. 6*, 89−98.
28. Zhou, A., Faint, R., Charlton, P., Dafforn, T. R., Carrell, R. W., and Lomas, D. A. (2001) *J. Biol. Chem. 276*, 9115−9122.
29. Quan, L. T., Caputo, A., Bleackley, R. C., Pickup, D. J., and Salvesen, G. S. (1995) *J. Biol. Chem. 270*, 10377−10379.
30. Lomas, D. A., Elliott, P. R., Chang, W. S., Wardell, M. R., and Carrell, R. W. (1995) *J. Biol. Chem. 270*, 5282−5288.
31. Xue, Y., Bjorquist, P., Inghardt, T., Linschoten, M., Musil, D., Sjolin, L., and Deinum, J. (1998) *Structure 6*, 627−636.
32. Bottomley, S. P., and Stone, S. R. (1998) *Protein Eng. 11*, 1243−1247.
33. Djie, M. Z., Stone, S. R., and Le Bonniec, B. F. (1997) *J. Biol. Chem. 272*, 16268−16273.
34. Calugaru, S. V., Swanson, R., and Olson, S. T. (2001) *J. Biol. Chem. 276*, 32446−32455.
35. Simonovic, M., Gettins, P. G. W., and Volz, K. (2000) *Protein Sci. 9*, 1423−1427.
36. Musil, D., Zucic, D., Turk, D., Engh, R. A., Mayr, I., Huber, R., Popovic, T., Turk, V., Towatari, T., and Katunuma, N. (1991) *EMBO J. 10*, 2321−2330.
37. Kamphuis, I. G., Kalk, K. H., Swarte, M. B., and Drenth, J. (1984) *J. Mol. Biol. 179*, 233−256.
38. Zhou, Q., Snipas, S., Orth, K., Muzio, M., Dixit, V. M., and Salvesen, G. S. (1997) *J. Biol. Chem. 272*, 7797−7800.
39. Bruice, P. Y. (1998) *Organic Chemistry*, 2nd ed., Prentice-Hall, Upper Saddle River, NJ.

BI0159985