Serpin Structure, Mechanism, and Function

Peter G. W. Gettins*

Department of Biochemistry & Molecular Genetics, University of Illinois at Chicago, M/C 536, 1819-53 West Polk Street, Chicago, Illinois 60612

Received February 8, 2002

Contents

* Correspondence address: Department of Biochemistry & Molecular Genetics, M/C 536, 1819-53 West Polk Street, Chicago, IL 60612. E-mail: pgettins@uic.edu.

1. Introduction

Serpins are a superfamily of proteins, whose membership is based on the presence of a single common core domain consisting of three β-sheets and 8−9 α-helices, and with a set of highly unusual structural and functional properties that result from the presence of this core domain (Figure 1). Many members are serine proteinase inhibitors, from which the family name derives, with a unique mechanism of action. The initial identification of this superfamily was based on an observation of primary structure similarities between the three proteins human antithrombin, human α_1-proteinase inhibitor, and chicken egg white ovalbumin,[1] which clearly established a relationship between these three proteins, despite a sequence identity of only 30% and seemingly very different biological functions, with the first two being inhibitors of serine proteinases and the last an abundant storage protein of egg white without proteinase inhibitory activity.[2] Only with the determi-

10.1021/cr010170+ CCC: $39.75
Published on Web 11/08/2002

Peter G. W. Gettins was born in Sunderland, UK, in 1953 and became a U.S. citizen in 1990. He received his M.A. in chemistry (1976) and D.Phil. in biochemistry (1979) from Oxford University. Work for his doctoral thesis, under the guidance of Raymond A. Dwek, was on the application of NMR to antibody combining sites. After postdoctoral work at Yale University under the supervision of Joseph E. Coleman, applying multinuclear NMR to the study of the metalloenzyme alkaline phosphatase, he took a position as Assistant Professor in the Department of Biochemistry at Vanderbilt University in 1984. In 1988, he was appointed Associate Professor and in 1993 moved to the Department of Biochemistry at the University of Illinois at Chicago as Professor. He is also currently the Interim Director of the Center for Structural Biology at the University of Illinois at Chicago. His research interests are in the structure and function of the superfamily of proteins known as serpins and of the low-density lipoprotein receptor-related protein (LRP) which binds and clears serpin-proteinase complexes and many other unrelated protein ligands. His laboratory uses a range of biophysical and biochemical approaches in these studies, including NMR and fluorescence spectroscopies, X-ray crystallography, calorimetry, and kinetic analysis.

nation of several crystal structures of serpins (see section 4) was it demonstrated that the primary structure similarity within a region of approximately 350 residues corresponds to a conserved tertiary structural domain. Although this core of about 350 residues is present in all serpins, with minor modifications due to small insertions or deletions that may introduce loops or lengthen or shorten α-helices, there is considerable variation in size among different serpins resulting from N- and C-terminal polypeptide extensions and, in some cases, N-type and/or O-type glycosylation. Thus, while many serpins have molecular weights of 40−60 kDa, C1 inhibitor, which has an ∼100 residue N-terminal extended tail-like extension that is heavily O- and N-glycosylated, migrates on SDS−PAGE as an ∼105 kDa protein.[3] This appears so far to be the largest known serpin, with 478 residues, six N-type carbohydrate chains, and 12 O-type carbohydrate chains. A recent search (in 2000) of all available protein sequences, using the sequence of human α_1-proteinase inhibitor as the query sequence, identified >400 serpins from all organisms.[4] With the completion of the sequence of the human genome as well as of genomes from a number of other organisms, many more serpins are likely to be identified. Despite the presence of an extensive, common fold in all serpins, the pairwise identity of primary structures can be as low as 25%.

2. Nomenclature and Gene Structure

The name serpin was originally coined in recognition of the fact that most serpins then identified were inhibitors of serine proteinases.[5] Even at that time, however, it was known that there were members of the superfamily that lacked any proteinase inhibitory properties, namely, angiotensinogen[6] and ovalbumin.[1] Since that time, not only has the number of noninhibitory serpins greatly increased, but serpins that inhibit cysteine proteinases have also been identified. Whereas the name serpin as an acronym describing the functional properties of the superfamily is therefore clearly inappropriate, it has established itself so well as a name for the structural superfamily that a committee charged by the HUGO Gene Nomenclature Committee with classifying members of the superfamily in a rational way recommended retention of the name for all superfamily members, with classification into clades based on phylogenetic relationships.[7]

Sixteen clades, designated A through P, have been identified so far, with an additional 10 serpins that are unclassified "orphans" (Table 1). Using this system, plasminogen activator inhibitor 1 is the first member of clade E and so is designated SERPINE1, with the gene symbol *SERPINE1*. Neither the alphabetic proximity of letters designating two clades nor the numerical designation of two serpins within a clade is meant to indicate evolutionary proximity. Newly discovered serpins will be classified using the existing clade designation where appropriate, or given a new clade letter identifier, using the next available letter. Within a clade, a new serpin will be given the next available numeric designation. Thus, with 10 human serpins already identified in clade A, the next member to be identified will be designated SERPINA11.

Table 1. Serpin Clades[a]

clade name	letter
α_1-proteinase inhibitor	A
intracellular, ov-serpin	B
antithrombin	C
heparin cofactor II	D
proteinase nexin, PAI-1	E
α_2-antiplasmin, PEDF	F
C1 inhibitor	G
HSP47	H
neuroserpin	I
horseshoe crab	J
insect	K
nematode	L
blood fluke	M
viral SPI1−2/CrmA-like	N
viral SPI3-like	O
plant	P
unclassifed (orphans)	

[a] From refs 4 and 7.

Human serpins have so far been found on 10 different chromosomes. However, 22 of the 34 identified are found in three clusters: four at 6p25, eight at 14q21.1, and 10 at 18q21.3. A smaller cluster of two is present at 3q26. Within each cluster, all serpins belong to the same clade. Furthermore, the clusters of genes on chromosomes 6 and 18 together belong to, and entirely constitute, clade B, the ovalbumin or ov-serpin clade,[8] and have evolved from a common ancestor by one or two interchromosomal

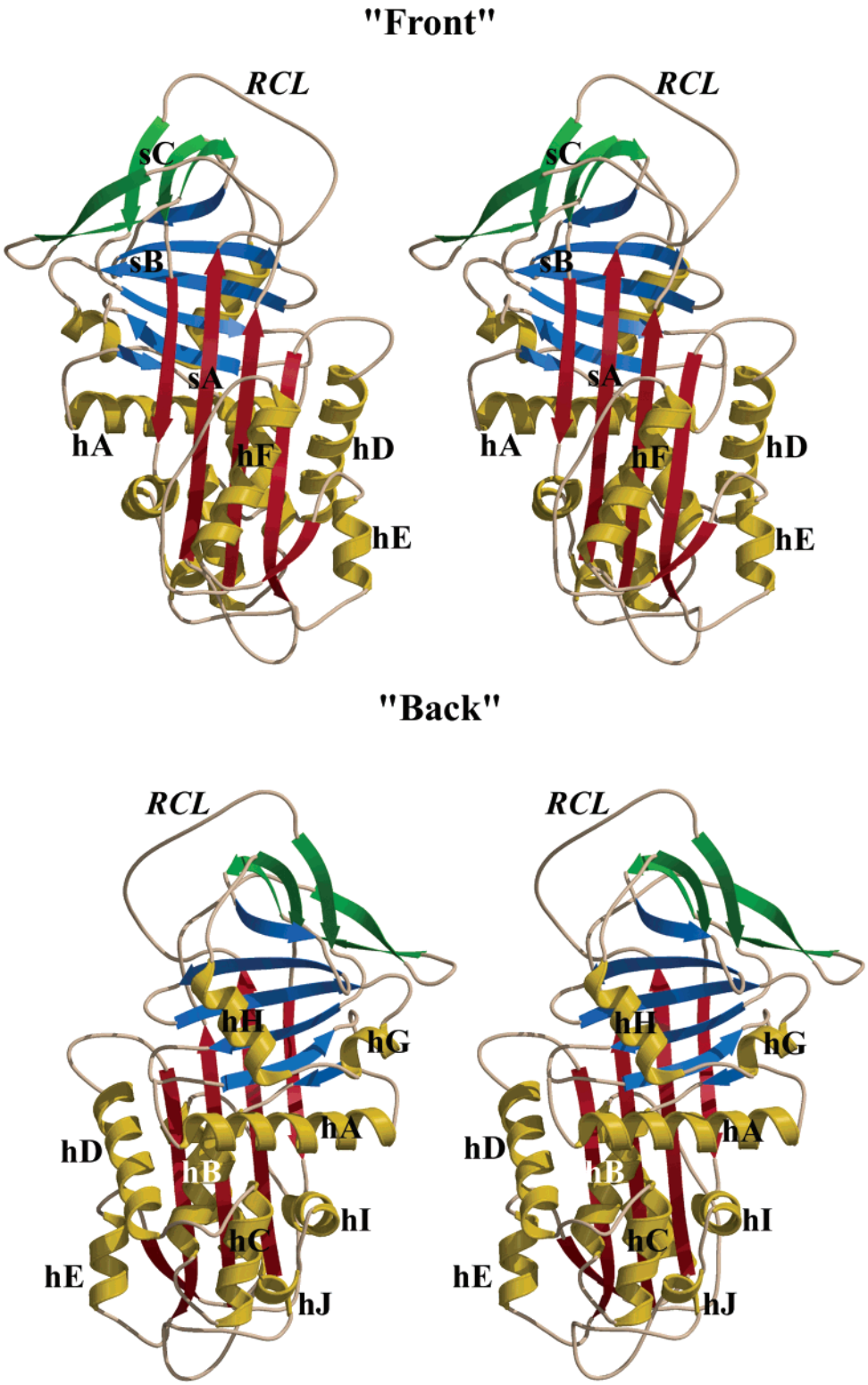

Figure 1. Front and back stereoviews of a typical serpin (here α_1-PI from the structure of Elliott et al.[98]) to illustrate the notable secondary structural features and the location of the reactive center loop. β-sheets A, B, and C are shown in red, blue, and green, respectively. The eight α-helices are designated A through H and are labeled. The reactive center loop (RCL) is indicated at the top of the molecule and rendered in yellow.

duplications and several intrachromosomal duplications.[9] With only two exceptions, all of the clade A serpins occur in the 14q21.1 cluster, which again has been proposed to have arisen by gene duplication. It has further been suggested that such gene duplication to produce clusters of serpins has occurred to produce serpins with properties related to the same main physiological function. Indeed, in the case of the intracellular ov-serpins, many of these have been shown to have properties related to inflammation and tumorigenesis. As examples, hurpin (SERPINB13) is underexpressed in squamous carcinomas of the oral cavity, but upregulated in psoriasis,[10,11] while maspin is a serpin identified by differential expression in

breast carcinoma cells, and shown to be an inhibitor of metastasis and a tumor suppressor,[12,13] a function that might be mediated by its ability to bind to type I and type III collagen.[14] Serpins B3 and B4 have as common names squamous cell carcinoma antigens 1 and 2.[15] The functional relationship of the cluster on chromosome 14 is perhaps less obvious, since it includes not only the inhibitory serpins α_1-PI and α_1-antichymotrypsin, but also the noninhibitory serpin corticosteroid binding globulin. In other organisms, there are also examples of clustering suggestive of evolution by gene duplication. Thus, in the tobacco hornworm, *Manduca sexta*, there is a remarkable alternative splicing pattern for serpin gene-1, such that exon 9, which encodes the reactive center loop, can be any of 12 different exons, giving rise to 12 serpins, each with the first 336 residues identical, and differences only in the C-terminal 39−46 residues.[16]

3. Occurrence and Function of Serpins

Serpins are extremely widely distributed among eukaryotes and in some viruses that infect them. They are absent from fungi and from chlorophytes, despite being found in higher plants.[4] Until very recently, it was also thought that they were absent from bacteria, though this has now been shown to be incorrect, with the first report of serpins also in prokaryotes.[17] While many serpins in multicellular organisms are extracellular, there are others that are intracellular (those of the ov-serpin sub-family, clade B[8]), and some that can exist in both extra- and intracellular forms.[18,19] Those that are intracellular have mostly been found within the cytoplasm, though there are now some reports of intracellular serpins that can be found within the nuclear compartment,[20,21] in a cell type-dependent manner. In this section, only an overview is given of the many serpins known. Where there are important mechanistic aspects of particular serpins, such as the rate acceleration of proteinase inhibition of antithrombin and heparin cofactor II by heparins, these are discussed in section 5. It should also be realized that the number of serpins known for a given species or genus is mostly a reflection of the scientific or commercial interest in that species rather than of the total serpin complement of the organisms. Thus, even the lowly nematode *Caenorhabditis elegans* is now known to have at least 8 serpins.

3.1 Human Serpins

Because of their importance in many highly regulated physiological processes, such as blood coagulation, fibrinolysis, and inflammation, and the consequent involvement in pathologies in cases of dysfunction,[22] human serpins are by far the best characterized as a group. So far, 34 serpins from nine clades have been identified and characterized to varying extents, mostly at the purified protein level[7] (Table 2).

Apart from the highly abundant pan-proteinase inhibitor α_2-macroglobulin, the most abundant human plasma proteinase inhibitors are nearly all serpins. Many of these have as a common target the proteinase thrombin. This is a proteinase that has procoagulant activity when targetting fibrinogen, and anticoagulant activity when bound to thrombomodulin, which changes its substrate specificity and makes it an effective activator of protein C.[23] Others are more specialized as inhibitors of the coagulation proteinases, such as factor Xa and XIa, or of the fibrinolytic proteinases plasmin, tPA and uPA. Several of these serpins (antithrombin, heparin cofactor II, PAI-1, proteinase nexin 1, and protein C inhibitor) are heparin-binding serpins, with consequent possible physiological consequences of protein localization and, for some proteinases, acceleration of the rate of inhibition. With the exception of protein C inhibitor, the heparin-binding serpins appear to use a structurally homologous region (helix D) as part of the positively charged heparin binding site, suggesting a common evolution, despite belonging to three different clades. Whereas knockouts of the gene have been carried out in mice for each of these proteins, only that for antithrombin is embryonically lethal, with death occurring by day 15.5 of gestation.[24]

Other extracellular inhibitory serpins with known important functions are α_1-PI, α_2-antiplasmin, C1 inhibitor, and neuroserpin. α_1-PI is abundant (1−2 mg mL^{-1}, but even higher levels occur during inflammation, since it is an acute phase protein) and is the principal inhibitor of elastase secreted by human neutrophils at sites of inflammation. The rate of this reaction, unaided by cofactors, is one of the fastest known for serpins (6.5×10^7 M^{-1} s^{-1}). Much work on this serpin has been done because of the high polymorphism it displays, with ~80 alleles identified,[25] and the relationship of such polymorphism to disease. In particular, α_1-PI deficiency has been linked to emphysema, a situation exacerbated for smokers as a result of oxidation of the P1 methionine to a sulfoxide. The sulfoxide form of the serpin reacts 2000-times less rapidly than the thioether form.[26] α_2-Antiplasmin is the principal inhibitor of plasmin, the proteinase responsible for cleavage of fibrin in blood clots. This is also a fast reaction, with second-order rate constant of 3.8×10^7 M^{-1} s^{-1}.[27] Despite the apparent importance of this serpin, a knockout is not embryonically lethal and the mouse shows enhanced fibrinolytic potential, but without overt bleeding in the unchallenged state.[28] C1 inhibitor is the only known inhibitor of the activated proteinases C1r and C1s that form part of the initiation complex of the classical complement pathway of the immune response. An unusual aspect of C1 inhibitor is that it is the largest human serpin known and that the extra length is accounted for by a very heavily glycosyated N-terminal tail of ~100 residues that contains both O-type and N-type carbohydrate. The role of this unique tail is not presently known. Although the precise role of neuroserpin is not known, there has been much recent interest in it in connection with its deposition as inclusion bodies in the brains of some familial dementia patients.[29] In vitro it has been shown to be an effective inhibitor of tPA.[30]

Of the 34 known human serpins, seven at least are noninhibitory, meaning that they are not known to

Table 2. Human Serpins, Using the Clade Designation[a]

serpin	synonyms	size	RCL (P4−P4′)	chromosomal location	target proteinases	rates of reaction (M^{-1} s^{-1})	comments	ref
				Clade A				
SERPINA1	α_1-proteinase inhibitor (α_1PI), α_1-antitrypsin (α_1AT)	394 (418)	AIPM- -SIPP	14q32.1	neutrophil elastase	6.5×10^7	Many allelic forms. Z variant associated with emphysems. Forms polymers.	26
SERPINA2	α_1-AT-related protein	399 (420)	EKAW- -SKYQ	14q32.1			Protein not characterized	402
SERPINA3	α_1-antichymotrypsin (α_1ACT)	398 (423)	ITLL- -SALV	14q32.1	cathepsin G chymase prostate specific antigen	5.1×10^7 4.3×10^4 8.9×10^1	Binds β-amyloid peptide. Complexes with PSA used in prostate cancer diagnosis	26, 414−416
SERPINA4	kallistatin, kallikrein inhibitor, PI4	401 (427)	IKFF- -SAQT	14q32.1	tissue kallikrein	1.4×10^4	In vivo function not established	221
SERPINA5	protein C inhibitor (PCI), plasminogen activator inhibitor-3 (PAI-3)	387 (406)	FTFR- -SARL	14q32.1	activated protein c uPA plasma kallikrein acrosin	2.5×10^4 9×10^4 1×10^5 5.6×10^7	Thrombin in complex with thrombomodulin inhibited rapidly	417−421
SERPINA6	corticosteroid binding globulin (CBG)	383 (405)	LNLT- -SKPI	14q32.1	elastase substrate		Noninhibitory, but RCL inserts upon cleavage. Corticosteroid transporter; hormone affinity reduced by cleavage.	31
SERPINA7	thyroxin binding globulin	395 (415)	LSDQ- -PENT	Xq22.2			Noninhibitory, but RCL inserts upon cleavage. Thyroxine transporter; hormone affinity reduced by cleavage.	31
SERPINA8	angiotensinogen (AGT)	452 (485)	NKPE- -VLEV	1q42-q43	renin substrate		Noninhibitory. Cleaved RCL does not insert. Renin cleaves in N-terminus at 10−11, gives angiotensin I peptide	422, 423
SERPINA9	centerin	434	FIVR- -SKDG	14q32.1			Involved in maturation of naive B cells	424
SERPINA10	protein Z-dependent protease inhibitor (ZPI)	423 (444)	ITAY- -SMPP	14q32.1	factor Xa, factor XIa		Requires protein Z as cofactor for Xa inhibition, but not XIa inhibition	290
				Clade B				
SERPINB1	leukocyte elastase inhibitor (LEI), monocyte/neutrophil elastase inhibitor (MNEI), PI2, EI, ELANH2	379	ATFC- -MLMP	6p25	neutrophil elastase cathepsin G proteinase 3	3.4×10^7 2.3×10^6 1.7×10^7	Broad range specificity for elastase-like and chymotrypsin-like proteinases. Uses adjacent P1 residues with different proteinases	425, 426
SERPINB2	plasminogen activator inhibitor-2 (PAI-2), placental PAI,	415	MTGR- -TGHG	18q21.3	uPA		Present in gingival inflammation in crevicular fluid.	427, 428
SERPINB3	squamous cell carcinoma antigen 1 (SCCA1)	390	GFGS- -SPAS	18q21.3	cathepsin K cathepsin L cathepsin S	1.4×10^5 3×10^5 5.2×10^5	Inhibitor of both cysteine and serine proteinases	303
SERPIN B4	squamous cell carcinoma antigen 2 (SCCA2), leupin	390	VVEL- -SSPS	18q21.3	cathepsin G chymase	1×10^5 1×10^4	Can be changed into cysteine proteinase inhibitor with small RCL changes	310, 429
SERPINB5	maspin, PI5	375	ILQH- -KDEL	18q21.3			Noninhibitory, inhibitor of metastasis, tumor suppressor. Binds to collagen types I and III	14, 15, 430
SERPINB6	placental thrombin inhibitor (PTI), PI6, cytoplasmic antiprotease (CAP)	376	MMMR- -CARF	6p25	thrombin plasmin chymotrypsin cathepsin G	1.2×10^5 2.3×10^6 6×10^6 6.8×10^6	May inhibit cathepsin G that leaks into cytoplasm in vivo	431−434

Table 2. Continued

serpin	synonyms	size	RCL (P4−P4′)	chromosomal location	target proteinases	rates of reaction (M^{-1} s^{-1})	comments	refs
SERPINB7	megsin	380	IVEK- -QLPQ	18q21.3			Found in glomerulus; up-regulated in IgA nephropathy	9, 435
SERPINB8	cytoplasmic antiprotease 2 (CAP2), PI8	374	RNSR- -CSRM	18q21.3	furin thrombin subtilisin A	6.5×10^5 1×10^5 1.2×10^6	Only ov-serpin shown to be inhibitor of subtilisin-like proteinase furin.	187, 436
SERPINB9	cytoplasmic antiprotease 3 (CAP3), PI9	376	VVAE- -CCME	6p25	granzyme B caspase 1 (ICE) subtilisin A	1.7×10^6 7×10^2 2.4×10^6	Inhibitor of proteinases of three different structural types.	186, 308, 437
SERPINB10	bomapin, PI10	397	IDIR- -IRVP	18q21.3	thrombin		Expression specific to bone marrow cells. CD loop contains nuclear localization sequence.	105, 438
SERPINB11	epipin	392	IAVK- -SLPM	18q21.3			Not yet characterized. NP_536723 gi	439
SERPINB12	yukopin	405	VSER- -SLRS	18q21.3	trypsin plasmin	2.5×10^6 1.6×10^4	Widely distributed. Does not inhibit thrombin, factor Xa or uPA.	439
SERPINB13	headpin, hurpin, PI13	391	FTVT- -SAPG	18q21.3			Target proteinases not yet known. Found in epidermal cells; overexpression in psoriasis	11, 440
SERPINBP1	PI8-like 1			6p25			probable pseudogene	
					Clade C			
SERPINC1	antithrombin (AT), antithrombin III (ATIII)	432 (464)	IAGR- -SLNP	1q23-q25	thrombin, factor Xa factor IXa	3.7×10^7 4.4×10^7 1×10^7	Requires heparin as cofactor for both Xa and thrombin reactions. Bridging and allosteric mechanisms	228, 242, 244
					Clade D			
SERPIND1	heparin cofactor II (HCII), leuserpin 2	480 (499)	FMPL- -STQV	22q11	thrombin	1.1×10^7	Requires dermatan sulfate or heparin-like GAG for fast inhibition. Allosteric mechanism.	441
					Clade E			
SERPINE1	plasminogen activator inhibitor-1 (PAI-1), endothelial PAI	379 (402)	VSAR- -MAPE	7q21.3-q22	uPA, tPA thrombin activated protein C	1.1×10^7 2.3×10^7 1.1×10^5 1.8×10^5	Vitronectin stabilizes active form. Inhibits cell migration and angiogenesis in vivo	285, 286, 442−444
SERPINE2	protease nexin 1 (PN1), glia derived nexin (GDN)	379 (398)	LIAR- -SSPP	2q33-q35	uPA, tPA thrombin acrosin factor XIa	1.5×10^5 3×10^4 1.2×10^9 2.4×10^6	Heparin acceleration for reactions with thrombin, acrosin, and factor XIa	154, 347, 445, 446
					Clade F			
SERPINF1	pigment epithelium-derived factor (PEDF), EPC-1	403 (418)	PAHL- -TFPL	17p13.3			Noninhibitory, non-RCL-insertable. Anti-angiogenic and neurotrophic factor	447, 448
SERPINF2	α_2-antiplasmin (α_2-AP), A2AP	452 (491)	AMSR- -MSLS	17pter-p12	plasmin	3.8×10^7	Inactivated by MMP-3	27, 296
					Clade G			
SERPING1	C1-inhibitor (C1 INH)	478 (500)	SVAR- -TLLV	11q11-q13.1	C1r, C1s plasma kallikrein	2.8×10^3 1.2×10^4 1.7×10^4	Only effective inhibitor of C1r and C1s	449−451

Table 2. Continued

serpin	synonyms	size	RCL (P4−P4′)	chromosomal location	target proteinases	rates of reaction (M^{-1} s^{-1})	comments	refs
				Clade H				
SERPINH1	heat shock protein 47 (HSP47), collagen-binding protein 1 (CBP1), colligin 1	400 (417)	EELR--SPKL	11p15			mRNA not found in human cells. May be a pseudogene.	49, 452, 453
SERPINH2	collagen binding protein 2 (CBP2), colligin 2	400 (418)	EELR--SPKL	11q13.5			Rheumatoid arthritis-related antigen. May be true human HSP47	454, 455
				Clade I				
SERPINI1	neuroserpin	394 (410)	AISR--MAVL	3q26	tPA uPA plasmin	1.5×10^5 4.7×10^4 1.1×10^5	Point mutants associated with polymeric deposits in the brain.	30, 456, 457
SERPINI2	myoepithelium-derived serine proteinase inhibitor, MEPI, pancpin	387 (405)	IPVI--MSLA	3q26			Pancreas-specific. Down-regulated in pancreatic cancer. Targets not identified.	458−460

[a] Adapted from Table B, of ref 7, Supplemental Material.

inhibit any proteinase (see 5.5.1). These are the hormone-binding serpins, corticosteroid binding globulin (CBG), and thyroxine binding globulin (TBG), the hormone precursor angiotensinogen (AGT), the tumor suppressor maspin, the anti-angiogenic factor pigment epithelium-derived factor (PEDF), and the two collagen chaperone proteins CBP-1 (HSP47, for heat shock protein 47) and CBP-2. Of these, only CBG and TBG appear to undergo spontaneous insertion of the reactive center loop into β-sheet A upon cleavage, with consequent reduction in hormone affinity.[31,32] For the others, spontaneous loop insertion does not appear to occur, presumably because of the presence in each case of reactive center loop residues that are inimical to such insertion and/or sheet expansion mechanisms that are slow to operate. The reason either for these proteins to be serpins or for the latter group to be designed not to undergo loop insertion is unclear. In each case, it appears that one or more regions outside of the reactive center loop are involved in their biological activities. Mouse knockouts of HSP47[33] and maspin are embryonically lethal, while that of angiotensinogen resulted in increased death during weaning.[34]

For many of the other human serpins, neither their specific roles nor physiological targets are well understood. Studies of rates of reaction and proteinase specificity are mostly guided by the residues in and around the P1−P1′ bond and are not necessarily indicative of the true in vivo targets. Here, it will be useful to know not only the rates of reaction of a given serpin with different possible target proteinases, but also the tissue distribution and local concentration in vivo of these components.[35]

3.2 Other Mammalian Serpins

Outside of humans, the best characterized occurrence of serpins in mammals, at least at the gene level, is in the mouse. Analogues of human α_1-PI and α_1-antichymotrypsin have been identified at loci termed *Spi-1* and *Spi-2*, respectively. Unlike the human genes, however, these loci represent clusters of proteins with hypervariability in the reactive center loop regions.[36] Another expansion of mouse analogues of human serpins occurs in clade B, the intracellular serpins. As with the human equivalents, these are found in clusters on two different chromosomes; 1[37] and 13[38] in mice compared with 6 and 18 in humans, but show much greater expansion in the mouse, probably reflecting a greater repertoire of proteinases encountered by the mouse.[39,40] Murine homologues of other human serpins have also been identified, including protein Z-dependent proteinase inhibitor[41] (clade A), antithrombin[42] (clade C), heparin cofactor II[43] (clade D), PAI-1[44] and protease nexin 1[45] (clade E), $\alpha_2$antiplasmin[28] and PEDF[46,47] (clade F), C1 inhibitor[48] (clade G), HSP47[49] (clade H), and neuroserpin[30] (clade I). Many other examples of individual mammalian serpins exist, including demonstration of equivalent gene clustering in the rat at the Spi loci.[50]

3.3. Plant Serpins

Relatively few plants have so far been shown to contain serpins, though this may again reflect the

narrow focus of the researchers involved in identifying them, who have tended to examine food grains. Accordingly, the best characterized are those from wheat, barley, and rye (clade P). In wheat, a set of five serpins has been cloned and characterized and shown to possess inhibitory activity and, for four of them, reactive center loops with sequences at the reactive bond resembling the repeated glutamine-rich sequences found in prolamin plant storage proteins, suggesting a possible role for these serpins in inhibiting insect proteinases that the insect might use to degrade the plant proteins.[51] Similar serpins have been found in rye.[52] Other serpins, with methionine and arginine P1 residues, have been found in barley[53,54] and shown to have proteinase inhibitory activity. In the pumpkin (*Cucurbita maxima*), a serpin has been identified in the soft tissue and shown to be an effective elastase inhibitor.[55] Again, a role in defense against insect predators has been suggested, in this case against sap-sucking insects. With the sequencing of the genome of *Arabidopsis thalia*, 13 serpins have also been identified there, but not characterized.

3.4 Viral Serpins

Although they are not survival factors for viral growth in culture, viral serpins are required for virulence and/or expression of the full range of infectivity. In the orthopoxviruses (variola, vaccinia, and rabbitpox), there are three serpins, designated SPI-1, SPI-2 (both clade N), and SPI-3 (clade O). Similarly in myxoma viruses there are also three serpins, designated SERP1 (clade E), SERP2 (orphan), and SERP3.[56] Of significant interest are the SPI-2-like and SERP2 serpins, which have a P1 aspartate and have as targets cysteine proteinases of the caspase family and the serine proteinase granzyme B. Through inhibition of these proteinases, the virus can abrogate host cell apoptosis and downregulate host inflammatory response. The mechanism of caspase inhibition by such serpins is discussed in more detail in section 5.8. SERP1 serpins, with P1 arginine, have been shown in vitro to be inhibitors of uPA, tPA, thrombin, and factor Xa.[57] Other viruses have also been shown to contain one or more serpins, including swinepox, lumpy skin disease virus, fowlpox,[58] and members of the *Rhadinovirus* genus.[7]

3.5 Drosophila and Other Arthropod Serpins

With the recent completion of the *Drosophila melanogaster* genome sequence, a much better understanding of the total serpin requirements of this organism will be forthcoming. It has been estimated that there are about 32 *Drosophila* genes coding for serpins. However, a search of GenBank for identified protein sequences so far yields only 13 complete serpin sequences. Reports characterizing individual *Drosophila* serpin proteins are also relatively scarce. Only one has been well characterized as a cloned protein.[59] Another has been identified as a gene encoding an active serpin,[60] and the presence of one has been inferred from its ability to form a higher molecular weight covalent complex with the serine proteinase Easter.[61] A study aimed at identifying proteins produced by the male accessory gland and secreted into seminal fluid identified 12 such proteins, one of which (Acp76A) was a serpin (orphan).[62] More recently, a serpin was identified as negatively regulating the antifungal defense mechanism of *Drosophila*, which is controlled by the *spaetzle/Toll/cactus* gene cassette, with the gene product spaetzle being proteolytically cleaved in the absence of the serpin and subsequently acting as the ligand for Toll.[63] This serpin, Spn43c, is one of a cluster of three that maps to the 43A1.2 locus.[64]

Serpins have also been identified in other insects, including notably the 12 *M. sexta* alternative splice products of the gene encoding serpin 1,[65] but also from the silkworm, *Bombyx mori*[66,67] (both clade K) and from the fall webworm moth, *Hyphantrea cunea* (so far an "orphan" serpin). Three coagulation inhibitors have been identified from the horseshoe crab (clade J),[68] *Limulus polyphemus*, and are together thought to control proteinases of the *Limulus* coagulation cascade.

Of some structural and mechanistic interest is another orphan serpin, AFXa,[69] from the mosquito *Aedes aegypti*. This serpin appears designed to be a potent *reversible* inhibitor of the host's factor Xa, and possesses a shortened reactive center loop, perhaps to facilitate it acting more like a classical protein proteinase inhibitor than as a serpin. This is discussed in more detail in section 5.9.

3.6 Nematode Serpins

The genome of *C. elegans* is another that has recently been sequenced and accordingly given insight into the full complement of serpins needed for functioning of a complete organism. An analysis of this genome directed at identifying serpins found eight sequences for complete serpin-like proteins (belonging to clade L), with six of these having reactive center loop sequences seemingly compatible with loop insertion and hence proteinase inhibition.[70] This relatively small number should be viewed in the context of eight chymotrypsin-like serine proteinases present in *C. elegans*. As yet, there are no reports of expression or characterization of these serpins.

3.7 Parasite Serpins

Five serpins have been identified in parasites, three that have related gene structures and so together constitute clade M, and two orphans, both from *Brugia malayi*. The related serpins are from *Schistosoma mansoni* (two examples) and *Schistosoma haematobium*. The latter is a 58 kDa glycoprotein expressed on the surface of the worms[71] and so may be involved in promoting evasion of the host's immune response. Individuals infected with this parasite generate antibodies to this serpin.[72] One of the *Brugia* serpins has been shown to be developmentally regulated, with transcription beginning between days 8 and 9 of larval development and continuing through adult and microfilarial stages.[73]

3.8 Avian Serpins

With few exceptions, examples of known avian serpins are restricted to chicken (*Gallus gallus*). Here seven serpins have been identified, three intracellular ones from clade B, and one each from clades C (antithrombin), D (heparin cofactor II),[74] H (hsp47),[75] and I (neuroserpin).[76] Of the intracellular serpins, the most studied is ovalbumin, one of the founding members of the serpin superfamily.[1] This is a noninhibitory, abundant serpin of unknown function, but usually viewed by default as a storage protein. Structural and mechanistically relevant aspects of ovalbumin are treated elsewhere in this review. The most interesting chicken serpin is MENT,[20] an intracellular serpin that possesses a large insertion between the C and D helices that contains a nuclear localization signal, a chromatin-binding domain, and an "A-T hook" binding motif. It is thought that it may promote chromatin condensation. Other avian serpins include ovalbumins from quail and turkey and antithrombin from ostrich.[77]

3.9 Fish Serpins

This is a relatively poorly studied genus, with examples of serpins restricted to two clade A, α_1-PI-like serpins from carp,[78−81] and antithrombins from salmon[82] and fugu fish. Nevertheless, the availability of antithrombin sequences from species so different from humans is of use in examining conserved features of heparin binding and activation.

3.10 Amphibian and Reptile Serpins

As with fish serpins, knowledge of the types and abundance of amphibian or reptile serpins is scant. The African clawed frog, *Xenopus laevis*, is best studied, with antithrombin,[77] heparin cofactor II,[74] α_1-PI and another clade A serpin EP-45[83] (for estrogen regulated protein) known. In addition, a recent sequence determination of turtle antithrombin[77] has been made.

4. Serpin Structures

X-ray and other structural studies of serpins and serpin complexes of various kinds have matured rapidly in recent years. Still by far, however, the greatest number of studies has been on uncomplexed serpins. These structures of noncomplexed forms of serpins are considered here, to illustrate both the common features of the serpin fold, and also the types of conformational change that can occur with this fold, which are critical to the functioning of serpins as a unique family of proteinase inhibitors. The more limited structural information on serpin complexes of various kinds is given in section 5, which deals with mechanism and regulation of activity.

The first structure of a serpin was obtained in 1984 and was that of cleaved human α_1-proteinase inhibitor.[84] It was a surprise in that, whereas native, functionally active α_1-PI together with the zymogen chymotrypsinogen were set up for crystallization, in an attempt to crystallize a complex of the serpin with a target proteinase, the structure had only the serpin present, in an inactive reactive center loop-cleaved form. Equally surprising was that the reactive center loop was no longer in an exposed location, but instead had been integrated into β-sheet A, the largest β-sheet, as the fourth strand. Evidently, the serpin had reacted with traces of active proteinase to give what we now know to be the covalent complex and this had, over time, dissociated to give the cleaved serpin, which then crystallized. It was six years before the next structure was determined, that of the native form of the noninhibitory serpin ovalbumin.[85] However, shortly afterward the structure of the reactive center loop-cleaved form of ovalbumin was solved, which showed a fundamental difference from that of the cleaved form of the inhibitory serpin α_1-PI, in that the cleaved reactive center loop had not inserted into β-sheet A.[86] This suggested a difference in behavior of inhibitory and noninhibitory serpins and provided a link between conformational change in serpins and their ability to inhibit proteinases. The next structure was that of the latent form of PAI-1,[87] which showed yet another variation in basic structure, in that the location of the reactive center loop of this single chain form was the same as for the cleaved form of the inhibitory α_1-PI, but with the uninserted part of the loop being composed of residues that otherwise would have been strand 1 of β-sheet C, here pulled away to provide a "return" to the top of the molecule. Only in 1994 were the first structures of uncleaved forms of inhibitory serpins reported: for antithrombin[88,89] and a variant of α_1-antichymotrypsin.[90] Since then more that 30 additional structures of uncomplexed serpins have been reported for a total of 12 different serpins (Table 3). These include an intracellular inhibitory serpin (PAI-2[91]), a noninhibitory but physiologically active serpin (PEDF[92]), and a viral serpin in cleaved form (crmA[93,94]).

4.1 The Native State

With the exception of antithrombin and heparin cofactor II, which show an unusual variation within β-sheet A discussed in section 5.7.2, the secondary and tertiary structures of all serpins show remarkable similarities within the core domain. The primary structure of α_1-PI and the location of the different secondary structural elements are given in Figure 1 as an example. β-Sheet A is the largest of the three β-sheets and is composed of five strands, the first a short strand of 5−6 residues and the remainder longer strands of similar length (12−15 residues) that span the long axis of the protein. With the exception of the relationship of the central strands to one another, the β-sheet is antiparallel. However, since insertion of the reactive center loop into β-sheet A occurs between the central parallel strands in an antiparallel manner, β-sheet A becomes fully antiparallel in the cleaved form of loop-inserted serpins. β-Sheets B and C are shorter and are composed of six and four strands, respectively. There are eight α-helices, designated A through H. Relative to the face of β-sheet A, all of these helices, except helix F, lie on the backside of the protein in the view given in Figure 2A. Helix F lies across the front of β-sheet

Table 3. X-ray Structures of Serpins and Their Complexes

serpin	structure	resolution (Å)	PDB #	year	ref
α_1-proteinase inhibitor (SERPINA1)	cleaved human plasma	3.0		1984	84
	cleaved human S variant (tetragonal I)	3.0	7API	1989	461
	cleaved human S variant (hexagonal)	3.0	8API	1989	461
	cleaved human S variant (tetragonal II)	3.0	9API	1989	461
	recombinant native. T59A, T68A, A70G	3.46	1KCT	1995	462
	recombinant native F51L	2.90	1PSI	1996	97
	recombinant native "multi 7" F51L, T59A, T68A, A70G, M374I, S381A, K387A	2.70	1ATU	1996	463
	recombinant native	2.00	1QLP	1999	98
	recombinant cleaved polymeric,	2.60	1QMB	1999	136
	recombinant P1 Arg cleaved polymeric	3.00	1D5S	2000	137
	recombinant native A70G	2.10	1HP7	2001	99
α_1-antichymotrypsin (SERPINA3)	cleaved human plasma	2.70	2ACH	1991	464
	native recombinant	2.50		1994	90
	cleaved recombinant, A347R (P12)	2.40	3CAA	1996	188
	cleaved recombinant, T349R (P14)	2.90	4CAA	1996	188
	cleaved recombinant A349R (P14)	2.10	1ASN	1998	465
	uncleaved recombinant L55P variant in δ-conformation	2.27	1QMN	2000	112
LEI (SERPINB1)	cleaved, horse leukocyte elastase inhibitor	1.95	1HLE	1992	466
PAI-2 (SERPINB2)	recombinant native, 68−98 deleted	2.0	1BY7	1999	91
	recombinant native, 68−98 deleted binary RCL peptide complex	1.60	1JRR	2001	195
ovalbumin	chicken egg, cleaved			1990	86
	chicken egg, native	1.95	1OVA	1991	467
	recombinant cleaved chicken, R339T, A352R	2.30	1JTI	2002	95
antithrombin (SERPINC1)	cleaved bovine plasma	3.0	1ATT	1993	468
	human plasma, heterodimer of cleaved and native molecules	3.2	1ATH	1994	88
	human plasma, heterodimer of latent and native molecules	3.0	1ANT	1994	89
	human plasma, heterodimer of latent and native molecules	2.6	2ANT	1997	101
	human plasma binary complex with high affinity heparin pentasaccharide	2.90	1AZX	1997	101
	recombinant human binary complex with RCL peptide	2.90	1BR8	1998	193
	human plasma α-form	2.90	1E05	2000	101
	human plasma β-form	2.60	1E04	2000	101
	human plasma α-form complex with high affinity heparin pentasaccharide	2.62	1E03	2000	101
	recombinant human N135Q, S380C (P14)	2.80	1DZG	2000	100
	recombinant human N135Q, S380C-fluorescein derivative	2.85	1DZH	2000	100
heparin cofactor II (SERPIND1)	native human	2.4	1JMJ	2002	104
PAI-1 (SERPINE1)	latent human plasma	2.60	1C5G	1991	87, 202
	cleaved human recombinant	2.70	9PAI	1995	469
	recombinant cleaved human A335E variant complexed with two RCL pentapeptides	1.95	1A7C	1998	194
	recombinant latent human	2.10	1DVN	1999	470
	recombinant native human	2.40	1DVM	1999	470
	recombinant native human	2.99	1B3K	1999	471
	recombinant native human	2.70	1BD2	2000	472
PEDF (SERPINF1)	recombinant native human	2.85	1IMV	2001	92
neuroserpin (SERPINI1)	recombinant mouse cleaved	3.06	1JJO	2001	473
serpin 1K (SERPINK)	recombinant *Manduca sexta* native	2.10	1SEK	1999	96
CrmA (SERPINN)	recombinant cleaved variant	2.90	1C8O	2000	94
	recombinant cleaved	2.50	1F0C	2000	93
protein−protein complexes	covalent, trypsin-α_1 proteinase inhibitor	2.60	1EZX	2000	139
	noncovalent, trypsin-*Manduca sexta* serpin 1K	2.30	1K9O	2001	160
	noncovalent, S195A thrombin-heparin cofactor II binary complex	2.2	1JMO	2002	104

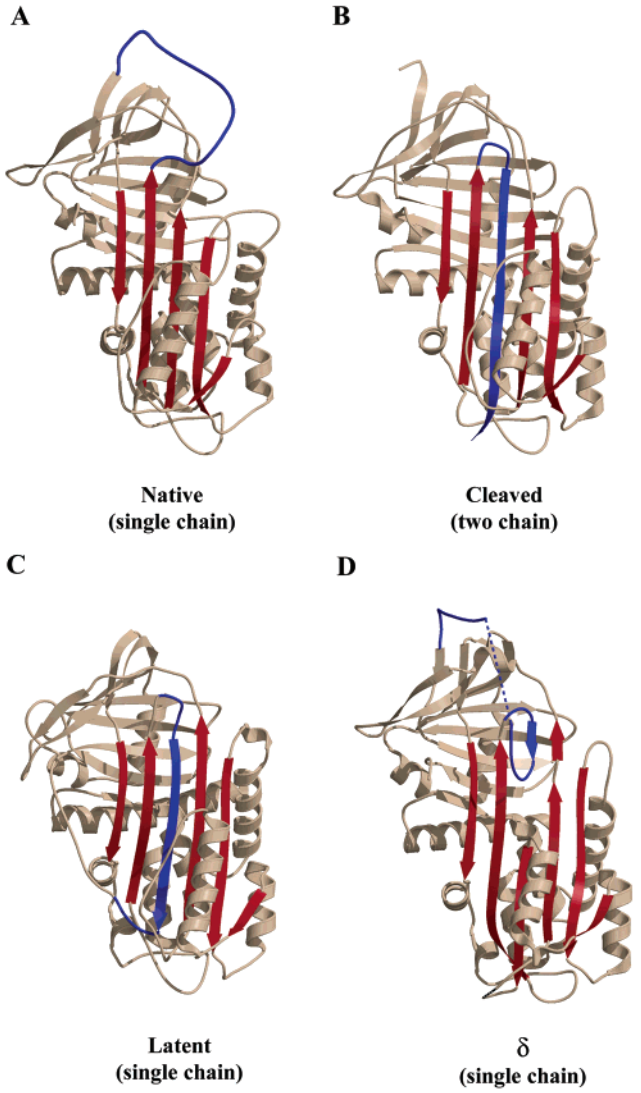

Figure 2. Comparison of the different conformational states open to serpins. (A) Native α_1PI (ref 98, pdb 1QLP), (B) cleaved α_1-PI (ref 461, pdb 7API), (C) latent PAI-1 (ref 202, pdb 1C5G), and (D) the δ conformation of a variant of α_1-antichymotrypsin (ref 112, pdb 1QMN). To emphasize the major changes in its location, the reactive center loop is shown in blue and the remainder of β-sheet A is in red.

A and may play a role in controlling the opening of the sheet or in stabilizing the five-strand conformation.[95] In the native conformation, the reactive center loop lies outside of the tertiary core of the serpin in a solvent-exposed environment, tethered at one end to the C-terminus of s3A and at the other to s1C. Within this loop lies the reactive bond that interacts with the active site of target proteinases. The length of the loop shows only small variation in inhibitory serpins, particularly on the side N-terminal to the reactive bond. Here the length is usually 17 residues, but is 16 for a few serpins. On the C-terminal side the length is less constrained and has been found to range from 5 to 9 residues in X-ray structures. The strict length requirement on the N-terminal side is directly related to the mechanism of inhibition and is discussed in detail below (see section 5.5.3). The conformation of the reactive center loop shows high variability, suggesting that there is no absolute structural requirement for function. In the case of native ovalbumin, the first native structure solved, the reactive center loop is a well-defined 3 turn α-helix.[85] In a recombinant variant of α_1-antichymotrypsin it is also helical, but closer to a 3_{10} helix.[90] However, in other serpin structures it has been found to either have no regular secondary structure[96] or in some cases to be in extended β-conformation.[97−99] A cautionary note, however, is that, since this region is invariably the most exposed and is often involved in protein−protein contacts in the crystal,[92,99] it may be the one region of the serpin for which the X-ray structure does not give a good representation of the behavior in solution. This is most likely to be the case with antithrombin, which has so far been crystallized only in heterodimeric forms composed of one molecule of latent, or cleaved, antithrombin and one of active serpin.[88,89,100,101] In the crystal structure residues P7 to P3 (387−391) of the reactive center loop of the active molecule replace the missing s1C of the latent molecule and consequently are in β-pleated conformation. What is not clear is whether the interaction with the latent molecule forces the reactive center loop of the active molecule into this conformation or whether its preference for this conformation in solution allows it to cocrystallize readily with a latent partner. Favoring the former is the finding that this conformation results in the P1 arginine residue pointing in toward the body of the serpin, where it can H-bond to glutamate 255. This would not be a good conformation to react with target proteinases, yet native antithrombin reacts rapidly with trypsin in the absence of heparin, which would require an accessible exposed arginine side chain.[102] A second unusual feature of the native antithrombin structures, and now also of native heparin cofactor II, is the partial expansion of β-sheet A at the top of the sheet caused by the insertion of the two residues at the far end of the reactive center loop (P15 and P14, corresponding to Gly382 and Ser383 in antithrombin and Gly430 and Thr431 in heparin cofactor II). This is based on X-ray structures[88,89] and fluorescence studies[103] of antithrombin and an X-ray stucture of heparin cofactor II.[104]

The viral serpin, crmA, is one of the smallest serpins and so its structure provides insight into the minimum structural features required for function.[93,94] Whereas the three β-sheets are fully present, there is a major shortening of helix A and a complete elimination of helix D. The greatly shortened region corresponding to this helix provides a short loop to β-sheet A and a novel short addition to β-sheet A as an additional strand (Figure 3). Large sequence differences between crmA and several otherwise nearly identical viral serpins in this region that forms the novel strand of β-sheet A have suggested a functional role for the region.[94] The ability to manipulate this N-terminal region, termed the C−D loop, without adversely affecting the ability of the serpin to undergo loop insertion and proteinase inhibition is in keeping with its location some distance from the interface with β-sheet A, and appears to be exploited in a number of serpins for insertions and deletions that may be connected with auxiliary properties. Thus, the C−D loop of the chicken protein

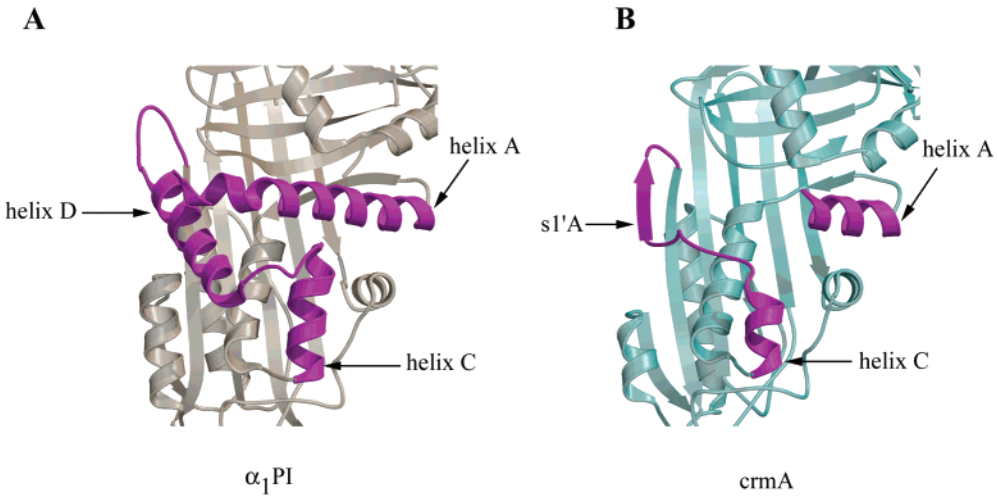

Figure 3. Closeup of the region of the C and D α-helices of α_1-PI (A) and their equivalents in the smaller viral serpin crmA (B), illustrating the ability to shorten this region by replacement of the D helix by a short, connecting strand without adversely affecting function. Conversely, in serpins such as chicken MENT[20] and bomapin[105] large, probably unstructured, insertions can be accommodated as surface loops between the C and D helices. Note also the dramatic shortening of the A helix in crmA.

MENT contains a 30-residue insertion that has little propensity for secondary structure formation and may thus form an extended "wing" capable of interacting with chromatin DNA, a property that may be key to its in vivo functioning.[20] In bomapin (SERPINB10) a 25-residue C−D insertion loop contains a highly positively charged nuclear localization domain that is essential for nuclear targetting.[105] In PAI-2, a similarly large insertion is required for its role in cell survival[106] and for transglutamination.[107] Although there is now an X-ray structure of PAI-2, it is with this region deleted. Nevertheless, the ability to fold such a truncated form of the protein and yet not disrupt other proteinase inhibitory properties attests to the auxiliary role of the insertion.

4.2 Conformational Change

The structures of cleaved forms of inhibitory serpins, determined first for α_1-PI and subsequently for other serpins, have all shown the same remarkable expansion of β-sheet A, through insertion of the cleaved reactive center loop as the fourth strand of the sheet, thereby perfecting the sheet as completely antiparallel, and changing the environment of the reactive center loop from completely solvent-exposed to mostly buried (Figure 2B). As a result of the insertion, the two ends of what had been the reactive bond have moved approximately 70 Å apart. Such loop insertion and expansion of β-sheet A also causes many additional smaller changes in the structure, resulting mainly from alterations in packing of underlying helices.[108] That this is an energetically favorable process is indicated by the change in unfolding temperature for a native serpin compared to a cleaved one. Whereas many native serpins unfold at ~60 °C, cleaved loop-inserted forms do so at >120 °C.[109] Since the transition from uncleaved to cleaved serpin is effectively irreversible, it is not experimentally feasible to determine the $\Delta G°$ for the process. However, calorimetric measurements of the heat released when the reactive center loop of a serpin is cleaved and inserts into β-sheet A show that $\Delta H°$ has a very large negative value (~50−60 kcal mol^{-1}).[110]

The third type of structure found for uncomplexed serpins is the latent form, first exemplified by PAI-1[87] (Figure 2C), and later seen in the heterodimeric structure of antithrombin.[89] In this structure, β-sheet A has undergone an expansion equivalent to that found in cleaved serpins, through insertion of the reactive center loop. The major difference from the cleaved form is that β-sheet C has given up the first strand, corresponding to the residues that are immediately C-terminal to the reactive bond, to provide a return from the bottom of β-sheet A back to β-sheet B. The much smaller increase in stability for the conversion of native to latent PAI-1 compared to the conversion of native to cleaved PAI-1[111] most probably results from the opposing energetic contributions from insertion of the reactive center loop into β-sheet A and the loss of stability from removing one strand of β-sheet C.

Finally, a fourth type of structure, designated δ, has been obtained for a naturally occurring variant of α_1-antichymotrypsin in which the reactive center loop has partially inserted into β-sheet A up to P12, with the place that would normally be occupied by P10−P3 being occupied by residues derived from the last turn of helix F and the loop connecting it to s3A[112] (Figure 2D). This is much more than the insertion of the hinge residues of the reactive center loop seen for active antithrombin, but much less than the full loop insertion of latent PAI-1. It has been taken as evidence that, rather than being an all-or-nothing phenomenon, loop insertion may proceed via defined intermediates, involving different extents of loop insertion.

Taken together, these four types of structure for monomeric serpins illustrate the types of conformational change that can occur as a result of having the serpin fold. The differences come about as a result of changes initiated by the reactive center loop and of the thermodynamic favorability of insertion of this loop into β-sheet A.

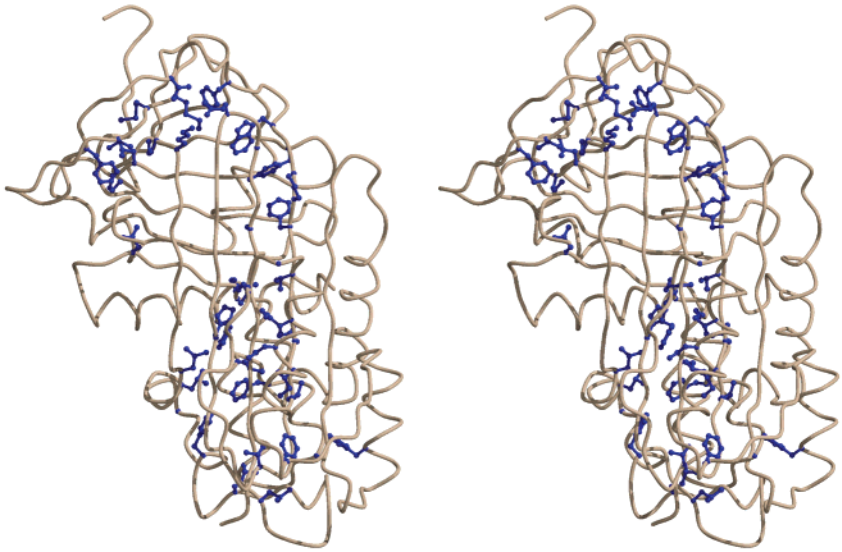

Figure 4. Location of the 51 core residues identified by Irving et al.[4] as being strictly conserved in >70% of serpins. Here shown, in stereo, within the structure of cleaved α_1-PI. All are core residues, with predominantly hydrophobic side chains needed for correct packing.

4.3 Serpin Metastability and Folding

Implicit in the recognition that the single chain form of a serpin can exist in a normal active form that has lower thermodynamic stability than the latent form is that serpins, as a family, are one of the few proteins that fold into a metastable state that is required for activity. The ability to routinely refold nondisulfide-containing serpins from guanidine-solubilized inclusion bodies into an active metastable state, indicates that, whatever the folding pathway to the metastable state, it must be the kinetically favored one. Relatively few studies have been carried out on the serpin folding pathway, so that it is still unclear what factors ensure that the metastable state folds faster than the latent state.[113−116] Since the major secondary structural differences between these states involve the reactive center loop, β-sheet A and β-sheet C, it seems plausible that the timing of formation and stability of one or more of these structural elements may be critical for directing the folding pathway. Given the variability in apparent reactive center loop structure and the frequent absence of stabilizing interactions once folded, it is more likely that formation of a four-strand β-sheet C and/or a five- rather than six-strand β-sheet A are the main factors rather than the reactive center loop.[117] Since the six-strand form of β-sheet A is completely antiparallel, whereas the five-strand form is not, and formation of the six-strand form, at least in cleaved serpins, is extremely favorable, it seems likely that it is early formation and stabilization of β-sheet C that directs the folding process toward the metastable state. Once the serpin is fully folded, the kinetic barrier to conversion to the latent form appears to be sufficiently high for most serpins,[118] with the notable exception of PAI-1, that the active metastable form is dominant under normal physiological conditions. In keeping with this idea, it was found that insertion of an extra 30 residues into the reactive center loop of α_1-PI circumvented the barrier to forming the most stable state, presumably by allowing both loop insertion *and* retention of strand 1 of β-sheet C to occur concurrently.[119] The resulting molecule, although not cleaved, had a stability toward guanidine unfolding similar to that of RCL-cleaved wild-type α_1-PI.

Sequence comparison of serpins within the region that corresponds to the α/β serpin core domain shows that about 51 residues have absolute invariance. These are mostly buried residues that are likely to be required for correct packing of the native state (Figure 4). The group of Yu in Korea has carried out extensive studies on increasing the stability of the native state by random mutagenesis.[120,121] This was done on α_1-PI with the result that seven main changes were found to give significant increases in stability, which were approximately additive.[122] This increase in stability was achieved without altering the ability of the protein to act as an efficient proteinase inhibitor, suggesting that the stabilizing mutations did not greatly hinder the ability of the molecule to undergo sheet expansion. Another random mutagenesis study was carried out on PAI-1 with a view to reducing the tendency to convert to the latent state. Up to a 100-fold increase in half-life for the spontaneous conversion was accomplished,[111,123] which approaches the normal half-life for such conversion of other serpins. These studies suggested that it is less favorable interactions of hydrophobic residues that stabilize the internal packing of the protein that are responsible for the decreased stability of PAI-1 relative to its latent state, compared with other serpins and that there is a functional advantage in PAI-1 for such packing defects that promote ease of conversion to the loop-inserted state.

4.4 Serpin Polymerization

The ability of serpins to fold into a metastable state, and to subsequently undergo insertion of the reactive loop into β-sheet A, is an essential part of the proteinase inhibition mechanism (see section 5). An unwanted consequence of this thermodynamically favored, conformational lability, however, is the tendency of serpins to use the same properties to form polymers through loop-sheet insertion mechanisms involving the loop of one molecule and a β-sheet of another. This can readily be seen with most serpins by heating at ~50−60 °C for a few minutes followed by analysis of the products on a nondenaturing gel.[124] A ladder pattern is typically seen that corresponds to noncovalent dimers, trimers, and higher order species. Such polymerization can also take place in vivo, particularly when there is a mutation that either affects the folding pathway to permit polymer formation or else alters the stability of the native state, such that even modestly elevated temperature can result in polymerization. The best characterized of such loop-sheet polymers is the Z-variant of α_1-PI. This is a frequent natural variant in some populations in which glutamate 342 has been changed to lysine. The phenotype of this mutation is a reduced circulating level of α_1-PI, and accumulation of sheetlike deposits of α_1-PI within hepatocytes[125] and also in the lung.[126] The reduction in circulating α_1-PI makes these individuals susceptible to emphysema.[126−131] The sheets are almost certainly extensive two-dimensional polymers that involve insertion of the reactive center loop of one molecule into β-sheet A of a second.[97,132] A possible structure of this is given in Figure 5A. Although it was at one time thought that α_1-PI Z variant polymerization resulted from reduced stability of the folded metastable state through loss of a stabilizing salt bridge interaction, it is now known to result from a folding defect in which a folding intermediate has a substantially increased lifetime.[133] This intermediate has a greater tendency to polymerize than the fully folded state. Another example of such polymeric deposits has been found with a brain-associated serpin, neuroserpin. Here a point mutation of serine 53 or serine 56 to proline or arginine, respectively, increases the tendency of the native state to undergo loop-sheet polymerization, with resulting deposition of polymers within the cerebral cortex.[29] A final example is an antithrombin variant, Rouen VI, in which fever (i.e., elevated temperature) led to thrombosis. The underlying basis for this seems to be that the mutation of Asn187 → Asp increases the tendency of the antithrombin to polymerize, especially at the elevated fever temperatures, and hence reduces circulating levels of functioning inhibitor.[134]

Though perhaps of lesser physiological importance, polymerization of serpins can also occur following nonproductive reactive center loop cleavage. This was first proposed by Mast and colleagues to explain the appearance of chainlike structures in electron microscope images of papain-cleaved α_1-PI.[135] Recent structure determinations by two groups of cleaved and polymerized forms of α_1-PI confirm the ability to form such polymers and show that they do involve head-to-tail insertion of the cleaved loop of one molecule into β-sheet A of another (Figure 5B).[136,137]

5. Mechanisms of Inhibition and Regulation

The most remarkable features of the serpin inhibitory mechanism are that, for inhibition of most proteinases, (i) there is a dramatic conformational change within the serpin as an essential part of distorting, and hence inhibiting, the proteinase, (ii) the inhibition represents kinetic trapping, and (iii) the trapped complex is covalent and effectively irreversible in nature. All of these features are in marked contrast to what occurs in almost all other classes of protein proteinase inhibitors, which instead use tight noncovalent association between the inhibitor and the proteinase, in a lock-and-key fashion, with little or no conformational change in either protein, to give a thermodynamically stable, but reversible complex. In the case of nonserpin serine proteinase inhibitors and proteinases, the nature of the binding interaction is extremely well understood from many X-ray structure determinations of a wide range of proteinase−proteinase inhibitor complexes, as well as of the individual component proteins.[138] In the case of serpin inhibition, there is very much less structural information, with only one X-ray structure of a covalent serpin−proteinase complex[139] and limited solution spectroscopic data (NMR and fluorescence) on similar complexes.[140−146] Both inhibition mechanisms and regulation of activity and function are considered in this section.

5.1 Inhibition by Non-Serpin Inhibitors

For nonserpin proteinase inhibitors the portion of the inhibitor reactive center loop that interacts with the proteinase adopts an extended β conformation, termed the "canonical conformation". This is true irrespective of the large differences in structure of the protein scaffolds (which determine their class) to which the reactive center loops are attached, and which hold these in place. The extent of interaction is variable, but always results from multiple contacts, that most commonly involve complementarity between the P6−P3′ region of the inhibitor and the S6−S3′ subsites of the proteinase.[147] Comparison of the structures of uncomplexed and complexed inhibitors always shows that there is little alteration in conformation necessary to bind optimally to the target proteinase, resulting in little or no loss of binding energy to modify the conformation.[138] The interaction is thus a good example of lock-and-key fit, which contrasts with the requirements for optimal substrate−proteinase interactions, which are better described as induced fits.[148] Such flexibility in substrates is required to permit adoption of a conformation in which the transition state is stabilized rather than the ground state, with resulting promotion of the proteolytic reaction. In contrast, the optimal interaction between the canonical conformation of nonserpin inhibitors and their cognate proteinases ensures that the ground state is stabilized relative to the transition state, with the result that the interaction results in inhibition. Since the catalytic

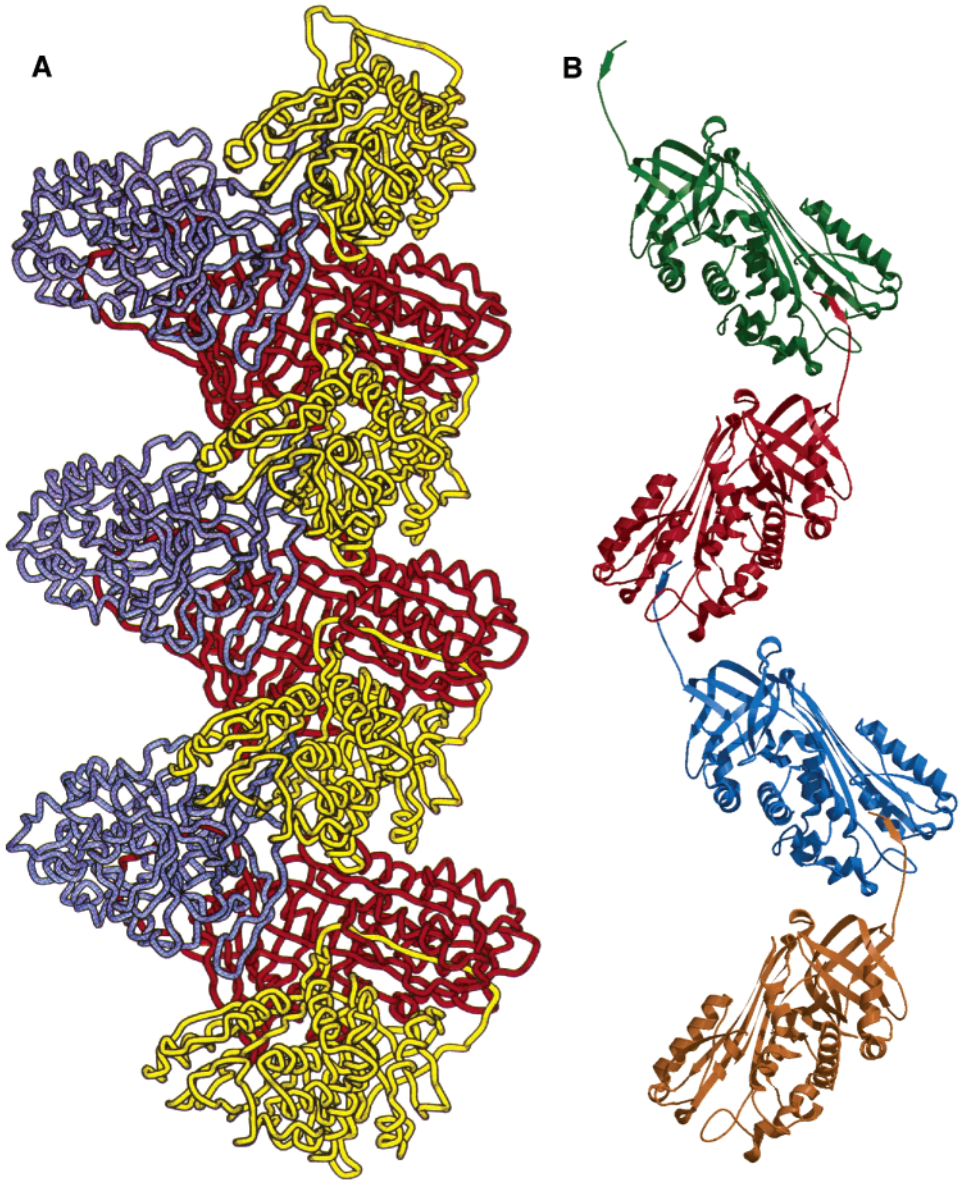

Figure 5. Two types of polymeric structure of serpins. Panel A, loop-sheet polymer structure suggested by Elliott et al.[97] to account for polymeric forms of Z variant α_1-PI, and involving insertion of the intact RCL of one molecule into β-sheet A of a second. Consecutive molecules are shown in different colors for ease of visualization. However, the conformation of each molecule is the same. Panel B, actual crystal structure determination[135] of a polymer formed from RCL-cleaved α_1-PI, in which the residual tail of the RCL, cleaved between P7 and P6, has inserted into the bottom of a second molecule of cleaved α_1-PI. Panel A reproduced from Elliott, P. R.; Lomas, D. A.; Carrell, R. W.; Abrahams, J. P. *Nat. Struct. Biol.* **1996**, *3*, 676−681. Copyright (1996) Macmillan Magazines, with permission.

serine is not involved in such interactions, it is not surprising that active site-modified proteinases, such as anhydrotrypsin, in which the active site serine has been modified to dehydroalanine, can bind to trypsin inhibitors with similar affinity to that of the active proteinase.[149]

5.2 The Serpin Branched Pathway Mechanism and Stoichiometry of Inhibition (SI)

The basic mechanism that can be applied to all serpin−serine proteinase inhibition reactions, and probably also serpin−cysteine proteinase inhibition reactions, is a branched pathway, suicide substrate inhibition mechanism that is outlined in Figure 6. In this, the proteinase recognizes the reactive center loop of the serpin, and in particular the reactive bond, as a potential *substrate*. The actions of the proteinase, both in initial recognition and in subsequent steps, are therefore those of a serine proteinase acting on a substrate peptide bond to cleave it. As such, the expected steps are (i) formation of an initial non-covalent Michaelis complex, (ii) attack of the active site serine on the peptide bond of the serpin to form a tetrahedral intermediate,[150] (iii) cleavage of the peptide bond to give a covalent acyl ester intermediate with release of the first product, the free amino group of the peptide bond, (iv) formation of the second tetrahedral intermediate through attack of water, and (v) departure of the second product.

In the scheme shown in Figure 6, the pathway as far as the branch point represents the normal substrate cleavage pathway of a serine proteinase, up

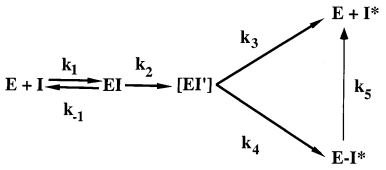

Figure 6. Branched pathway mechanism of serpins as suicide substrate inhibitors, showing only the essential intermediates. I represents the serpin and E represents the proteinase. The initial encounter, with forward rate constant of k_1, and back rate constant of k_{-1}, is the reversible formation of the noncovalent Michaelis-like complex, EI. This complex progresses through the normal covalent tetrahedral intermediate of a substrate cleavage reaction by a serine (or cysteine) proteinase to the acyl enzyme intermediate EI′, with overall rate constant of k_2. Only at this point, with the first leaving group departed and the RCL cleaved, can the RCL begin to insert into β-sheet A, translocating the proteinase and committing the intermediate to kinetic trapping as the complex EI*****, with rate constant k_4. Since there are continuing opportunities during translocation for the proteinase to complete the substrate reaction, with rate constant k_3, and escape, leaving behind cleaved serpin, I*****, the branch point between the substrate and complex pathways must be considered poorly defined in structural terms, and dependent on the particular serpin−proteinase pair and the relative values of k_3 and k_4. The kinetically trapped proteinase in EI* may decay to cleaved serpin and free proteinase with rate constant k_5, where k_5 is usually $\ll k_3$.

to and including initial cleavage of the peptide bond and concomitant formation of the covalent acyl ester intermediate between the active site serine of the proteinase and the carboxyl of the P1 residue of the serpin reactive center loop. From here onward there are two competing pathways that lead to distinct outcomes. The first pathway is continuation of the proteolysis reaction and subsequent release of a cleaved form of the serpin. This occurs with overall rate constant of k_3. The second pathway is trapping of the acyl intermediate by compromising the effectiveness of the proteinase to complete the proteolysis reaction, as a result of the loop-insertion conformational change within the serpin and the consequent distortion of the proteinase active site. This occurs with an overall rate constant of k_4. As discussed below (section 5.5.4), catalytic compromise of the proteinase may be a progressive process that increases as the reactive center loop inserts. Once this inactivation has occurred to the extent that completion of the loop insertion process can proceed without the possibility of deacylation occurring in that time, the acyl intermediate may be considered committed to that branch of the pathway. This branch represents the "suicide substrate inhibition" in that the proteinase has been inactivated as a result of its reaction with the serpin substrate. The effectiveness with which the proteinase is trapped is reflected in the rate constant for dissociation of the complex, k_5, since this represents the rate constant for completion of the proteolysis reaction, but from an acyl intermediate with a proteinase of compromised function. It has been estimated that k_5 is, at a minimum, 5−7 orders of magnitude smaller than k_3.[151]

Steps leading up to formation of the acyl enzyme intermediate are typically rate limiting, so that the values of k_3 and k_4 do not affect the overall rate of reaction of the serpin and proteinase. However, their relative values do determine the relative proportions of the acyl enzyme complex that are either stabilized as trapped covalent complex or continue to cleaved serpin, with release of active proteinase. If $k_3 \ll k_4$, the reaction is predominantly an inhibitory one, whereas if $k_3 \gg k_4$, the reaction is predominantly a substrate reaction with little proteinase successfully trapped as covalent complex. This balance between substrate and inhibition reactions leads to the useful concept of a "stoichiometry of inhibition" or SI, which is defined as the ratio of mols of serpin needed to inhibit 1 mol of proteinase. In terms of the rate constants of the pathway (and assuming that k_5 is small enough to be ignored, which is usually the case) SI is given by eq 1.

$$\mathrm{SI} = (k_3 + k_4)/k_4 \qquad (1)$$

Thus, where $k_3 \ll k_4$ and the reaction is nearly completely inhibitory, SI ∼1, whereas when $k_3 \gg k_4$ and the reaction is mostly a substrate reaction, SI ≫ 1. Since the acyl enzyme intermediate that serpins form with serine−proteinases is usually SDS-stable, an estimate of SI can be obtained from an SDS−polyacrylamide gel of reaction products. If the cleaved serpin can be distinguished from the native species on such a gel, the relative intensities of the bands corresponding to the high molecular weight covalent complex (corrected for the staining contribution of the proteinase moiety) and to the cleaved form are equal to k_4/k_3, from which SI can be calculated. This is true as long as the complex itself is kinetically stable prior to denaturation and that denaturation does not somehow promote deacylation and so perturb the product distribution present beforehand. In addition, given the increased proteolytic susceptibility of proteinases in complex with serpin, there may be more than one band representing covalent complex that needs to be quantitated.

Most kinetic assays of serpin inhibition, whether it is a discontinuous assay in which serpin and proteinase are incubated for different times and the reaction mixture is then assayed for residual proteolytic activity, or a continuous assay carried out in the presence of a chromogenic or fluorigenic substrate and monitoring the reduction in rate of substrate hydrolysis as a reflection of inhibition of the proteinase by the serpin, give an apparent overall association rate constant for covalent complex formation, k_{apparent}, that is an underestimate of the true rate of association and that results from recycling of proteinase from that fraction of the acyl enzyme complex that follows the substrate branch of the pathway. When the serpin/substrate pair has SI ∼ 1 the error is very small, but for reactions where k_3 is similar in magnitude to k_4, or greater, this will lead to larger errors. The value of k_{apparent} under these conditions is given by eq 2 (again under conditions where k_5 is too small to be a significant contributor to regeneration of free proteinase).

$$k_{\mathrm{apparent}} = k_2/K_{\mathrm{M}} \times 1/\mathrm{SI} \qquad (2)$$

Table 4. Kinetic Properties and SIs for Selected Recombinant Serpin Variants

variant	proteinase	SI	k_{app}	$k_{app} \times$ SI (k_2/K_S)	$k_{apparent}$ (w.t.)	ref
		P14				
antithrombin S → W	thrombin	10	3×10^2	3×10^3	8.5×10^3	10^3
	trypsin	40	5×10^3	2×10^5	1.3×10^5	10^3
antithrombin S → E	trypsin	195[a]	2.9×10^4	5.8×10^6	1.9×10^6	255
	thrombin	650[a]	1.1×10^4	7.2×10^6	1.7×10^7	255
	factor xa	550[a]	6.9×10^3	3.8×10^6	2.7×10^6	255
α_1-PI T → R	porcine elastase	70	1.6×10^3	1.1×10^5	2.8×10^5	152
		P12				
α_1-PI A → T	HNE	1.0	1.0×10^7	1.0×10^7	1.2×10^7	192
	trypsin	1.2	5.4×10^5	6.5×10^5	2.7×10^5	192
α_1-PI A → E	HNE	4.2	1×10^6	4.2×10^6	1.2×10^7	201
		P11				
α_1-PI A → E	HNE	1.5	0.7×10^6	1×10^7	1.2×10^7	201
		P10				
α_1-PI G → P	HNE	160	6.2×10^4	1×10^7	1.2×10^7	192
	trypsin	5.6	1×10^5	5.4×10^5	2.7×10^5	192
α_1-PI G → E	HNE	1.0	1.4×10^7	1.4×10^7	1.2×10^7	201
		P5				
α_1-PI E → G	HNE	3.8	1.2×10^6	4.5×10^6	1.2×10^7	474
	trypsin	2.1	3.4×10^4	7.2×10^4	1.9×10^5	474
		P1				
α_1-antichymotrypsin L → F	chymase	7.0	0.2×10^5	1.7×10^5	1.1×10^5	414

[a] Measured in the presence of heparin. In the absence of heparin SIs are much lower.

To obtain the true measure of the second-order rate constant for association of serpin and proteinase to form the prebranch point complex (i.e., k_2/K_M), it is therefore necessary to multiply $k_{apparent}$ by SI. This gives the correct overall rate constant for formation of the acyl enzyme intermediate and hence of steps involved in recognition of the serpin by the proteinase and of conversion to the critical acyl intermediate.

This rate constant (k_2/K_M) is very useful for evaluating the effects of mutations or other perturbations on different steps of the pathway, since it provides a way of separating effects that influence steps preceding acyl enzyme formation from those that follow and that determine only the product distribution. Thus, a mutation within the serpin might not affect the recognition by proteinase or the rate of formation of the acyl intermediate, yet lead to an increase in SI due to a slowing of k_4. If only $k_{apparent}$ were determined, the effect of the mutation would appear to be a reduction in the rate constant for the reaction. When corrected for SI, however, the true effect on the reaction rate (no perturbation) is obtained. At the same time, the fact that SI has increased reveals that the mutation has affected the relative values of k_3 and k_4. Other mutations might affect both the rate of acyl complex formation and the SI. The effects of mutations for a number of serpin/proteinase reactions have been examined in this way (Table 4). These include P14, P12, and P10 variants of α_1-PI, a number of reactive center loop variants of antithrombin, including mutations both in the hinge region and closer to the reactive bond, and a P14 variant of α_1-antichymotrypsin and provide examples of different kinds of effects.

The ability of the branched pathway mechanism to satisfactorily explain the effects of mutations on rates of reaction of serpin−proteinase pairs and on changes in SI was historically a major argument in favor of such a mechanism rather than of parallel, independent pathways leading to complexed or cleaved serpin. Thus, if parallel independent pathways existed, a mutation that did not affect the substrate pathway (k_3 unaffected), but reduced the rate of the inhibitory pathway (k_4 reduced), would result not only in an increase in SI, but also a reduction in the overall rate of consumption of the serpin. This could potentially be a very large reduction if the SI for the wild-type serpin were close to 1, since it would imply $k_4 \gg k_3$. This is not, however, what is observed. Instead, an unaltered rate of serpin consumption, as predicted by the branched pathway mechanism, is seen for such mutations.[152]

As explained below (see section 5.5.1), the structural basis for the increases in SI that are caused by many mutations in the hinge region has to do with the effect on the rate at which the reactive center loop can insert into β-sheet A. However, for a given serpin/proteinase reaction there are other factors that can also influence the relative rates of reactive center loop insertion and substrate cleavage and hence affect SI. These include alteration in temperature, pH, ionic strength, or of a cofactor that differentially affects reacted and unreacted species (Table 5). Thus, the reaction of C1-inhibitor with α-kallikrein, kallikrein light chain, plasmin and C1s showed changes in SI for a given reaction pair upon changing the temperature from 4 to 38 °C.[153] In each case, the SI fell upon increasing the temperature, as a result of an increasing fraction of inhibition at the higher temperature. Thus, while both k_3 and k_4 increased as a result of increase in temperature, the relative increase of k_4 was greater, resulting, in these examples, from a slower rate constant and hence higher activation energy for the inhibitory branch. A given increase in temperature results in a higher percentage effect on the reaction with the higher activation energy. Analogous temperature-dependent changes have been seen for the reaction of protease nexin 1

Table 5. Effects of Changes in Environment on SI Values for Different Serpin−Proteinase Pairs

serpin−proteinase pair	normal SI	environmental change	effect on SI	ref
C1-inhibitor−C1s	1.05 at 38 °C	reduction in temperature to 4 °C	SI increases to 1.2	153
C1-inhibitor−kallikrein	2.5 at 38 °C	reduction in temperature to 4 °C	SI increases to 6.3	153
protease nexin 1−acrosin	37 at 37 °C	reduction in temperature to 12 °C	SI increases to 3000	154
α_1-PI−subtilisin Carlsberg	8 at 37 °C	reduction in temperature to 0 °C	SI increases to 23	185
α_1-PI−proteinase K	5 at 37 °C	reduction in temperature to 0 °C	SI increases to 13	185
α_1-antichymotrypsin−proteinase K	9 at 37 °C	reduction in temperature to 0 °C	SI increases to ∼70	185
α_1-antichymotrypsin−chymase	5 at pH 8	reduction in pH to 6	SI decreases to 1.3	475
α_1-antichymotrypsin−chymase	5 at pH 8 and I 0.15	increase in ionic strength	SI decreases	475
α_1-antichymotrypsin−chymotrypsin (A349R variant)	5 in the presence of Cl^-	change of buffer anion to F^-	SI increases to 15	475
PAI-1 (P14T → H)−uPA	∼1 at pH 8.0	reduction in pH to 5	SI increases to 40	476
antithrombin−thrombin/heparin	1.0 at I 0.15	reduction in ionic strength to 0.01	SI increases to 9.8	155

with acrosin, which has an SI of 37 at 37 °C, but an SI of ∼3000 when the temperature is reduced to 12 °C.[154] For the reaction of antithrombin with thrombin catalyzed by heparin, the effect of reducing the ionic strength from 0.15 to 0.01 is to increase the SI from 1.5 to 10.[155] This occurs as a result of tighter heparin binding to the proteinase and a slowing of the rate of insertion of the reactive center loop and the tethered proteinase into β-sheet A.

An important point to realize from the effects of various environmental factors on SI is that serpins, unlike nonserpin proteinase inhibitors, are thus capable of having their reactions regulated, not only in the overall rate of inhibition but in the *outcome* of the reaction, inhibited proteinase−serpin complex vs cleaved serpin. While it might seem wasteful to cause an increase in SI, the increased production of cleaved serpin may itself act as a regulatory signal or else provide a species with desirable properties. Examples may be the antiangiogenic activity ascribed to cleaved antithrombin,[156] and of chemotactic activity associated with cleaved forms of both α_1-proteinase inhibitor[157] and α_1-antichymotrypsin[158] (see section 5.7.5).

5.3 The Initial Noncovalent Michaelis-like Complex

A critical step in the ultimate kinetic trapping of a proteinase by a serpin is the initial recognition step that leads to formation of the noncovalent Michaelis-like complex. In nonserpin inhibitors, this is the *only* step involved in inhibition and so all factors that influence specificity and stability of the complex depend on the nature and extent of interactions between the two proteins within the complex. Because serpin inhibition involves reaction beyond formation of the initial noncovalent complex, it is only aspects of specificity and rate of reaction that are likely to be influenced by the nature of the noncovalent complex.

Light has recently been shed on the initial interactions of serpin and proteinase by two structural studies on noncovalent complexes between a serpin and an inactivated serine proteinase. The first was a solution 2D NMR study of the complex between S195A trypsin and α_1-PI−Pittsburgh (P1 Met → Arg),[159] while the second was an X-ray crystal structure determination of the complex between *M. sexta* serpin 1K and S195A trypsin.[160] In the NMR study, 2D [^{1}H−^{15}N] HSQC spectra of ^{15}N alanine-labeled component proteins allowed examination of the conformation of both serpin and proteinase. The serpin spectrum, in particular, was extremely sensitive to conformational changes, such as those caused by insertion of the reactive center loop into β-sheet A, as judged by significant shifts of *all* alanine resonances between the native and cleaved conformations. It was found that the conformation of the body of the serpin in the complex was identical to that of the native inhibitor, and that the only perturbation of any alanine residues were of those at P4 (residue 355) and alanine 258, which lies directly underneath P4 and is in contact with it. No loop insertion of any reactive center loop residue was detected, and the alanines in the remainder of the reactive center loop (P12, P11, and P9) had unperturbed chemical shifts, indicating identical local structure of the reactive center loop. In the trypsin moiety, complex formation caused no change in the fold of the proteinase, with the only alanines perturbed being those five that are directly in, or adjacent to, the active site (195A, 55A, and 56A) or at the bottom of the S1 specificity pocket. These five alanines are also the only ones perturbed in noncovalent complexes with BPTI and SBTI (as seen by equivalent NMR spectra of their complexes), for which X-ray structures also confirm no change in tertiary structure for the body of the proteinase.[161] Somewhat surprisingly, the NMR studies also showed that the reactive center loop, which has considerable freedom of movement relative to the body of the serpin in the native state, retained much of this mobility in the noncovalent complex. These results indicated that the noncovalent complex involved relatively few contacts between the proteinase and the serpin and that these involved only the reactive center loop. The contacts presumably involved the P1 residue, given its importance in determining interaction with arginine-specific proteinases, and extended as far as P4, but did not involve residues closer to the hinge point. No conformational change occurred in either the serpin body or in the proteinase. Finally, the resulting complex was not rigid, but instead involved independent motion of the proteinase and interacting reactive center loop relative to the body of the serpin.

The X-ray crystal structure of the complex of the same proteinase with a different serpin (Table 3) agreed with the NMR study in some respects but differed in others. Thus, in agreement with the NMR study, no insertion of the reactive center loop into β-sheet A was found and the body of the serpin was

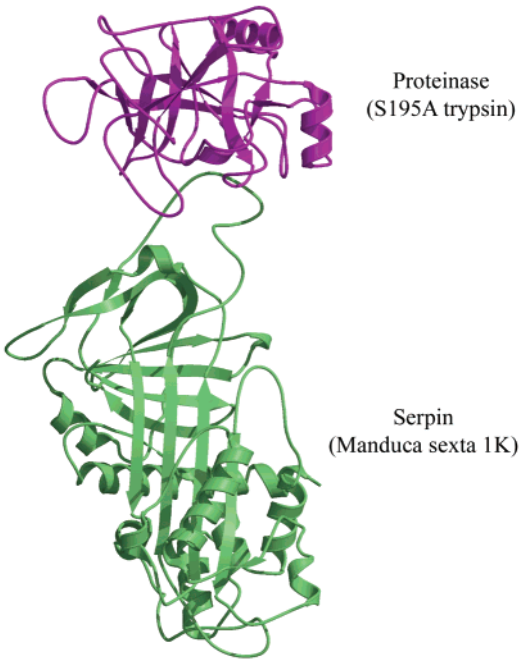

Figure 7. X-ray crystal structure of the noncovalent complex of *Manduca sexta* serpin 1K and S195A trypsin (PDB file 1K9O).[160] The body of the serpin can be superimposed on that of native serpin 1K (PDB file 1SEK)[96]. Similarly, the trypsin body can be superimposed on that of uncomplexed trypsin. There are extensive contacts between the reactive center loop, between P4 and P4′, and subsites on the surface of the proteinase. The serpin is rendered in green and trypsin in violet.

almost completely superimposable on that of the native protein, with rms deviation outside of the reactive center loop of 1.2 Å (Figure 7). Similarly, complexed trypsin had minimal conformational changes relative to free trypsin, with rms deviation of 0.9 Å. The P1 lysine side chain was, as expected, found in the S1 specificity pocket of the proteinase. The differences in conclusions between the NMR study and the X-ray study lay principally in changes in the reactive center loop. Whereas the NMR study showed no perturbation of the hinge region alanines, and by implication no contacts with the proteinase or structural changes much beyond P4, the X-ray structure found contacts as far as the P10 and P9 residues, and suggested a much more extensive contact region than in trypsin-BPTI. In addition, some changes in a distant part of the serpin (the F1 helix and connecting loop) were found, despite the absence of structural changes in the whole intervening region. No comment was made on the possible independence of the trypsin-reactive center loop region from the body of the serpin. The very recent X-ray structure of a noncovalent complex formed between S195A thrombin and heparin cofactor II also shows contacts between the proteinase and the serpin body,[104] suggesting that the particulars of the interaction between serpin and proteinase may depend on the pair and be influenced by whether they represents a cognate, physiologically relevant pair.

While the differences between the solution NMR study and the X-ray crystal study may reflect real structural differences in the two complexes that arise from the use of different serpins in each case, they may also reflect the difficulty of making conclusions about the reactive center loop of serpins from crystal structure determinations. As discussed above in considering the structures of native serpins, the greatest degree of variability in serpin structures from one determination to another is in the reactive center loop region. This is true even for the same serpin, which may result from the ability of this exposed flexible region to easily adopt different structures in different crystal forms, especially where contact interactions with other asymmetric units occur, or where different crystallization conditions are used. In addition, the requirement for high lattice order to obtain a diffracting crystal largely precludes study of the true dynamics that would occur in solution. In the particular case in question, an NMR investigation of the same complex could certainly determine whether there are more extensive contacts between the reactive center loop and the proteinase than were found with α_1-proteinase inhibitor Pittsburgh and trypsin.

In regard to this question of solution versus crystal studies, a study of fluorescently labeled PAI-1 species, with fluorophore introduced covalently at either P1′ or P9 positions, showed that, while both fluorophores reported binding of both anhydrotrypsin and S195A tPA, the fluorophore at P9 caused no alteration in binding affinity or kinetics, whereas that at P1′ had a 10−60-fold effect, suggesting that in these two systems, the extent of reactive center loop interaction between the serpin and either proteinase did not extend to the P9 position.[162] This same study also showed that the strength of the interaction between anhydroproteinase and PAI-1was not decreased by P14 mutation or by annealing a reactive center loop peptide into β-sheet A, indicating that no loop insertion occurred in the noncovalent complex, and thus supporting the conclusion of both the NMR and X-ray studies. This consensus makes the results of another fluorescence study very suspect. In that study the same saturation kinetics were found for *covalent* complex formation between α_1-proteinase inhibitor and elastase as for *noncovalent* complex formation between anhydroelastase and the same serpin. This was interpreted to indicate both that the proteinase had been translocated by reactive center loop insertion in the noncovalent complex and that translocation occurred prior to formation of the acyl enzyme intermediate.[163]

5.4 The Covalent Complex

A hallmark of the inhibition of serine proteinases by serpins is that the complex has a mobility on SDS−PAGE that is lower than that of either the proteinase or the serpin, and instead corresponds to that expected from a species containing both proteinase and serpin components. Early characterization of such high molecular weight species showed that they involved a covalent ester linkage between the serine proteinase active site serine 195 γO and the carboxyl of the reactive bond of the serpin,[164−166] and that they could be dissociated into cleaved serpin and

active proteinase by strong nucleophiles,[167,168] suggesting that the SDS-stable complex involves the same kind of acyl ester linkage that would be found as part of the normal cleavage pathway of a serine proteinase. Despite such findings there was, for a long time, considerable skepticism that such an intermediate existed in solution under physiological conditions. Instead, it was proposed that the acyl enzyme complex was an artifact of SDS-denaturation and that the complex present in solution was either a tetrahedral intermediate between the proteinase serine γO and the serpin reactive bond,[169] or even a noncovalent complex in which partial insertion of the reactive center loop into β-sheet A induced the remaining exposed loop into a canonical conformation with which the proteinase bound through a classical lock-and-key type interaction.[170] Nevertheless, evidence accumulated in favor both of the need for the active site serine for complex formation,[171,172] and of the complex involving cleavage of the peptide bond and formation of an acyl linkage.[173] In keeping with such acyl enzyme intermediate formation and the need for loop insertion as part of the inhibition mechanism, it was proposed that the structure of the irreversible serpin−proteinase complex consisted of a serpin moiety in which the cleaved reactive center loop has fully inserted into β-sheet A as strand 4, and of a covalently bound proteinase moiety that has consequently been translocated to the distal end of β-sheet A.[174,175] Fortunately, the controversy has now been largely resolved by direct structural studies on such complexes.

The first direct evidence that the proteinase does indeed move significantly upon formation of the complex was from fluorescence resonance energy transfer measurements on a complex between trypsin and α_1-PI Pittsburgh.[140] The extent of such movement was subsequently shown to involve complete translocation of the proteinase by over 70 Å,[142−144] as had been hypothesized (Figure 8A). While capable of accurately placing the proteinase and serpin components relative to one another in the complex, these fluorescence studies could not, however, give any direct structural information on the conformation of either the serpin or the proteinase. Such information was finally provided in 2000 with the determination by X-ray crystallography of the first serpin-proteinase complex (Table 3), that of bovine trypsin with recombinant α_1-proteinase inhibitor[139] (Figure 8B).

The X-ray structure confirmed that the inhibited complex (i) involved end-to-end translocation of the proteinase from the initial docking site, (ii) involved full insertion of the cleaved reactive center loop into β-sheet A, (iii) was a covalent species involving an acyl ester linkage to the γO of the proteinase active site serine, and (iv) involved a major disruption of the proteinase active site, with removal of the P1 side chain from the S1 specificity pocket, loss of the oxyanion hole that is required in the hydrolysis mechanism for stabilization of the tetrahedral intermediates and movement of the catalytic serine by 6 Å from its partner histidine 57. A surprising finding was that about 40% of the trypsin structure was crystallographically disordered and therefore could not be defined in the final model.This X-ray structure nicely explained much additional data in the literature. Thus, a study that showed that the P1 side chain in the covalent complex of a P1W variant of α_1-proteinase inhibitor with chymotrypsin was in a different inaccessible environment from that in the noncovalent complex[176] is understandable in terms of the side chain in the covalent complex still being within the active site of the proteinase and therefore sequestered, but no longer in the S1 pocket of the proteinase, whereas it would have been in this pocket in the noncovalent complex.[160] NMR studies which had shown a change in pK_a of the active site histidine of the complexed proteinase,[177] are likewise understandable in terms of an altered interaction with the active site serine. Even an antibody epitope mapping study[178] that, at the time it was published, seemed to be at odds with full loop insertion for the covalent complex, is easily understood in light of the actual X-ray structure. Thus, the monoclonal antibody recognized a four-residue three-dimensional epitope at the base of helix F together with one residue in the nearby loop connecting the C-terminal end of helix F with s3A and was capable of binding to a PAI-1/tPA complex as effectively as to native PAI-1. In the structures of covalent complex of α_1-PI/trypsin derived from X-ray crystallography[139] and by FRET,[143] the residues that are structurally equivalent to those of PAI-1 are solvent exposed and therefore, in PAI-1, should be accessible to antibody. It had been incorrectly assumed by the authors that a "bottom" placement of the proteinase would obscure this site, in part because the location imagined for such a bottom placement differed from that actually found, where the proteinase is moved partly to the side. Most importantly, the structure explained why the acyl enzyme intermediate is hydrolyzed extremely slowly, since it revealed radical alterations in the active site of critical features that are needed for efficient catalysis. These are the loss of the oxyanion hole and movement of the Ser195 γO away from the catalytic triad histidine 57. While there might be concern that the X-ray structure, which derives from analysis of crystals grown over a two-week period, might differ from that present within the first few minutes of reaction of serpin and proteinase, and thus from the physiologically relevant one, it is unlikely to be a valid concern. Thus, the fluorescence studies that came to the same structural conclusion as the X-ray study showed that the end point of fluorescence change was rapidly attained and was stable for many minutes afterward.[140,179,180]

One question about the complex that is still unresolved by the crystal structure is what has happened to the global conformation of the proteinase beyond the active site. The 60% of the protein whose electron density is well defined showed similar structure to native trypsin, though with high-temperature factors in some regions. For the ~40% of the trypsin that was crystallographically disordered nothing could be said of the final structure (or structures). A 2D NMR study of an almost identical complex (trypsin with the P1 Pittsburgh variant of α_1-PI) found that, whereas the serpin moiety of the complex was con-

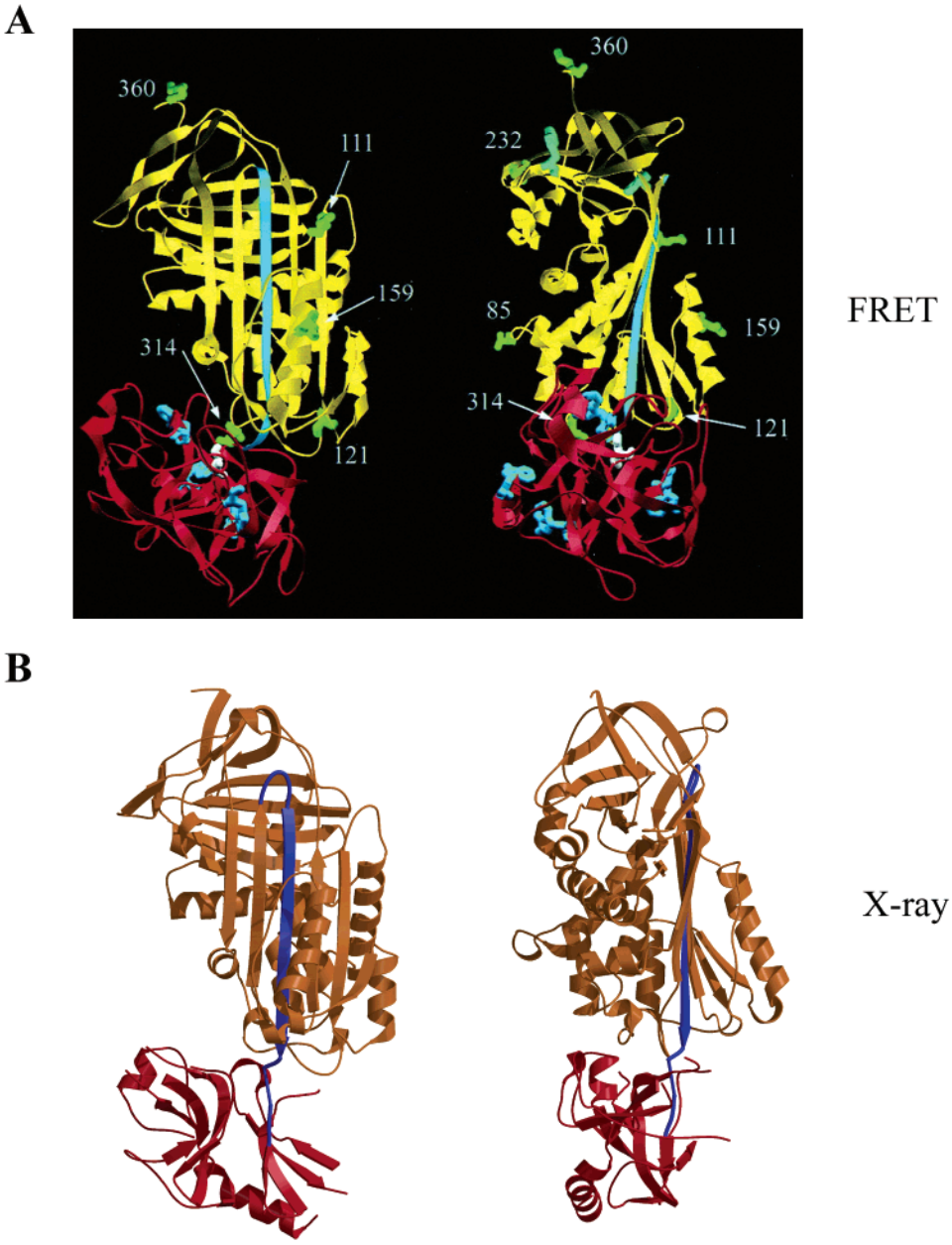

Figure 8. Comparison of the structures of the covalent complex between α_1-PI and trypsin determined (A) by fluorescence resonance energy transfer distance constraints[143] and (B) by X-ray crystallography.[139] Orthogonal views are shown for each structure. The FRET structure was modeled using trypsin and cleaved α_1-PI as rigid bodies, with distortion of loops in the contact region to allow close enough approach of the proteinase active site serine 195 γO to the P1 carbonyl of the inserted reactive center loop. The serpin is shown in yellow, the inserted reactive center loop is in blue and trypsin is in red. Numbered residues on the serpin are sites of fluorescent labeling used in the FRET distance determinations. The X-ray structure has a slightly different placement of the trypsin and is missing electron density of about 40% of the proteinase, and so appears smaller that in the FRET structure. The proteinase is again shown in red, the reactive center loop is in blue and the serpin is in orange. FRET figure reproduced from Stratikos, E.; Gettins, P. G. W. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 4808−4813. Copyright (1999) National Academy of Sciences.

formationally well defined and essentially identical to cleaved serpin (as in the X-ray study), the proteinase gave a spectrum with only few well dispersed and resolved resonances, suggesting that conformational interconversion was taking place in solution for a significant part of the protein.[145] While this study thus also failed to define the structure of the body of the proteinase in the complex, it did lend support to the idea that it is conformationally labile and that the X-ray crystal structure reflects what is actually occurring in solution. Another aspect of this conformational lability is the well-documented increase in proteolytic sensitivity of certain regions of proteinases when in complex with serpins.[109,181−183] These sites of enhanced proteolytic susceptibility map to the portion of the proteinase that is undefined in the X-ray crystal structure.

Since there is only one X-ray crystal structure so far of a covalent serpin−proteinase complex, there are still questions of how general such a structure is and hence how general such a mechanism of inhibition might be. Two fluorescence studies on PAI-1 and one on chymotrypsin/α_1-antichymotrypsin address this question. The first study on PAI-1, using donor−

donor energy transfer gave a model for the covalent complex between PAI-1 and urokinase-type plasminogen activator that also involves analogous full loop insertion and thus proteinase translocation.[144] The second on PAI-1 used contact perturbations of exogenous fluorophores to show similar structures for covalent complexes of PAI-1 with a series of arginine-specific proteinases: trypsin, low molecular weight and high molecular weight urokinase, and tissue-type plasminogen activator.[146] Together with the structures of the trypsin complex with α_1-PI, these studies show the same structural organization for the complexes of the same proteinase (trypsin) with two serpins (α_1-PI and PAI-1) and of four proteinases with one serpin, lending support to the generality of such a mechanism of inhibition for irreversible serpin−proteinase complexes. In contrast, the third study, on chymotrypsin/α_1-antichymotrypsin, came to the remarkable conclusion that the covalent complex involved essentially no movement of the proteinase from its initial interaction site.[184] It was proposed that this might be a feature unique to α_1-antichymotrypsin, made possible by the ability of (a variant of) α_1-antichymotrypsin to adopt a partially loop-inserted δ-conformation, as described in section 4[112] and shown in Figure 2d. It was suggested that in this conformation there might be stabilizing interactions between the proteinase and the reactive center loop. Quite against such a proposition is that the δ conformation is inactive as an inhibitor, suggesting inaccessibility of the P1−P1′ bond to the proteinase, rather than a set of stabilizing interactions. In addition, such a conformation offers no basis for stabilization of the acyl intermediate against continued hydroylsis.

Finally, there is the question of whether the mechanism of inhibition used by serpins against serine proteinases that have the chymotrypsin-like fold is the same as for inhibition of subtilisin-fold serine proteinases. Examples of such proteinases are furin, subtilisin A, proteinase K, and subtilisin Carlsberg. In terms of operation of a branched pathway mechanism and formation of SDS-stable complexes that represent kinetically trapped intermediates, this has been directly addressed by a study of α_1-PI and α_1-antichymotrypsin reactions with proteinase K and subtilisin Carlsberg.[185] Although SI values were in all cases high, the same mechanism of inhibition appeared to operate as for chymotrypsin-like proteinases, with α_1-PI capable of inhibiting subtilisin Carlsberg and α_1-antichymotrypsin capable of inhibiting proteinase K. It has also been shown that PI8 (SERPINB8) is an effective inhibitor of both furin and subtilisin A, and operates by the normal serpin mechanism.[186,187] Given the physical mechanism of kinetic trapping of chymotrypsin-like serine proteinases discussed above, it should not be surprising that serine proteinases with different folds should be similarly capable of active site distortion through compression.

5.5 Formation and Stability of the Serpin–Proteinase Complex

The description of the covalent serpin-proteinase complex given above requires that the reactive center loop insert into β-sheet A, carrying the proteinase with it and, that upon reaching its final resting place, the proteinase be distorted in such a way as to compromise its catalytic efficiency to a sufficient extent that it is effectively inhibited in a kinetic trap at the acyl intermediate stage of the hydrolysis reaction. In this section I consider those factors that affect the achievement of the final loop-inserted complex state, as well as those that affect the stability of this state once it has been achieved.

5.5.1 Requirement for Rapid Loop Insertion

The branched pathway mechanism of serpins as suicide substrate inhibitors requires that, for efficient inhibition (i.e., SI close to 1), the inhibitory branch of the pathway must operate much faster than the normal rate of hydrolysis of the acyl enzyme intermediate (i.e., $k_4 \gg k_3$). Taken together with the structure of the serpin−proteinase complex, which demonstrates that inhibition results from distortion of the proteinase active site, and which in turn results from the full insertion of the reactive center loop which presses the proteinase against the bottom of the serpin, this requires that the rate of insertion of the reactive center loop relative to the substrate cleavage reaction will determine the efficiency of the inhibitory pathway. The insertion involves incorporation of a previously exposed loop region into β-sheet A. As a result, alternating reactive center loop residues that previously had a solvent-exposed environment, become buried in the interior of the protein. The composition of the loop, particularly the residues that become buried, is therefore of great importance in influencing the rate of loop insertion. In addition, since the P14 residue is the first side chain to become buried, it is not surprising that it plays an especially important role.

The first demonstration of the importance of P14 was an examination of the SIs of a P14 Thr → Arg (residue 345) variant of α_1-proteinase inhibitor against several proteinases.[152] Whereas the SIs of the wild-type serpin against human neutrophil elastase, porcine pancreatic elastase or trypsin are all ~1, those of the T345R variant were increased to 45, 70, and 9.5, respectively. Although absolute rate constants for each branch of the pathway were not measured, it is unlikely that the changes in SI are due to change in the rate of the substrate reactions, since the mutation was 13 residues removed from the reactive peptide bond and so unlikely to affect the rate of such substrate reactions. The greatly increased SI values therefore require large reductions in the rates of the inhibitory branch, arising from a reduction in rate of insertion of the reactive center loop. Since the P14 side chain would normally become buried upon reactive center loop insertion, the replacement of the small, neutral threonine by a charged arginine would be much less favorable and probably increase the activation energy for insertion. Indeed, an X-ray structure of the cleaved form of an analogous P14 arginine variant of α_1-antichymotrypsin showed that to accommodate a charged residue the side chain of the arginine was turned outward rather then being buried, which then prevented the P14 backbone from

Table 6. Hinge Region Variants of Serpins with Substrate-like Behavior

serpin	variant	mutation	consequence	disease	ref
antithrombin	Hamilton	P12 (res. 382) Ala → Thr	decreased inhibition	thrombosis	203
antithrombin	Cambridge I	P10 (res 384) Ala → Pro	decreased inhibition	thrombosis	477
antithrombin	Cambridge II	P10 (res 384) Ala → Ser	decreased inhibition	thrombosis	204
C1 inhibitor	We	P14 (res 432) Val → Glu	no inhibition, substrate	angioedema	478
C1 inhibitor	Ma	P12 (res 434) Ala → Glu	no inhibition, substrate	angioedema	479, 480
C1 inhibitor	Mo	P10 (res 436) Ala → Thr	no inhibition	angioedema	481, 482
C1 inhibitor		P10 (res 436) Ala → Val	no inhibition	angioedema	480

forming hydrogen bonds with its flanking strands,[188] though subsequent residues were still able to form normal H-bonds with flanking strands of the sheet, with the expected register.

The most detailed examination of the effects of P14 mutations on the rate of loop insertion, and in turn on SI, has been recently reported for the reaction of PAI-1 with three proteinases. Replacement of the P14 threonine with 16 of the 19 possible amino acids gave variants that could all still form SDS-stable complexes with thrombin, uPA or tPA, but with SI values that were increased to greater or lesser extents depending on the nature of the mutation.[189] The largest increases in SI were found for the charged amino acids, which also caused the largest reductions in the rate constants for reactive center loop insertion, as determined using an attached P9 fluorophore as a reporter group. Of these, arginine caused the greatest effect, reducing the rate constant for insertion from $\geq 3\ s^{-1}$ for wild-type PAI-1 to $0.001\ s^{-1}$.

The complement of this type of study has also been carried out on a noninhibitory serpin, ovalbumin, with the goal of converting it into an inhibitor by increasing the rate of reactive center loop insertion. The X-ray structures of native[85] and cleaved[86] ovalbumin had shown that, whereas cleavage of the reactive center loops of inhibitory serpins resulted in spontaneous insertion of the reactive center loop into β-sheet A, the reactive center loop of ovalbumin did not insert at all. That this was a kinetic rather than a thermodynamic effect, however, had been suggested by a study on a conformational isomer of ovalbumin, S-ovalbumin, (a species that can be generated by incubation at elevated temperatures), which has properties consistent with partial insertion of the reactive center loop into β-sheet A.[190] Extending these studies by site-directed mutagenesis, it was found that mutating the P14 arginine of ovalbumin to serine, with or without changes at P12 and P11, to make the sequence more like inhibitory serpins, resulted in spontaneous insertion of the reactive center loop upon cleavage, up to at least position P9.[191] Indeed full loop insertion can occur, as shown by the recent X-ray structure of a slightly different ovalbumin variant, if sufficient time is allowed and there is no danger of shortening of the temporarily exposed loop through proteolysis.[95]

Other P-even residues (e.g., P12 and P10) can also have large effects on the rate of loop insertion and thus on the SI of the reaction. This was first noted from the partial or complete loss of inhibitory behavior of a number of naturally occurring variants of two serpins (Table 6) (though see below concerning the possible exceptional status of antithrombin). On the basis of such observations, the effects of hinge region mutations in α_1-proteinase inhibitor were more systematically examined, with changes being introduced at P12, from alanine to threonine and at P10 from glycine to proline.[192] Whereas the relatively conservative change at P12 had almost no measurable effect on SI for reaction with either HNE or trypsin, the more drastic change to proline at P10 increased SI to 5.6 for reaction with trypsin and to >100 for reaction with HNE.

Thus, whereas mutations at P14, P12, and P10 can each have major effects on the rate of loop insertion, it seems that the most critical residue is probably P14, the first side chain to insert. In wild-type serpins, once this residue has inserted it is likely that subsequent residues insert with increasing ease, at least until the blocking F-helix is encountered at about P9. There are two lines of evidence in support of this, involving annealing or displacement of reactive center loop peptides.

It has been shown that peptides corresponding to the 14 residues P14−P1 of a serpin can anneal with the native serpin by a mechanism analogous to the loop insertion that occurs upon cleavage of the reactive center loop.[193−195] Such complexes are inactive as inhibitors due to the inability of the reactive center loop to rapidly displace the tetradecapeptide;[196,197] evidence that was in fact used to support the requirement of loop insertion for proteinase inhibition. It was found that, whereas complexes of α_1-PI with even shorter peptides that contain the P14−P8 region still lacked inhibitory properties, complexes with peptides as long as a dodecapeptide that were missing residues P14−P13 retained inhibitory activity,[198] presumably reflecting the ease of displacement by the reactive center loop as it inserts. More recently, a study that examined the rate of formation of binary peptide complexes with the serpin PAI-2 found that the wild-type P14 residue needed to be present to bring about ready peptide incorporation, and if absent or mutated to valine, no peptide was incorporated under the conditions of the experiment.[199] In contrast, the P13 residue could be mutated from glutamate to glutamine or even lysine without affecting the rate of peptide annealing. Both types of study thus indicate the critical need for P14 in promoting loop insertion and in stabilizing the inserted conformation. Indeed, a recent study on PAI-2 examined the interactions of the P14 threonine side chain with internal residues in the crystal structure of the PAI-2−peptide complex and showed that a hydrogen-bonding interaction with tyrosine 258, that could also be formed in many other serpins where such a threonine/serine:tyrosine pair, is highly conserved.[195]

Table 7. Reactive Center Loop Sequences for Inhibitory and Noninhibitory Human Serpins[a]

serpin	P15	P14	P13	P12	P11	P10	P9	P8	P7	P6	P5	P4	P3	P2	P1	P1′	P2′	P3′	P4′
Inhibitory																			
A1; α_1-PI	G	T	E	A	A	G	A	M	F	L	E	A	I	P	M	S	I	P	P
A2	G	T	E	A	T	G	A	P	H	L	E	E	K	A	W	S	K	Y	Q
A3; α_1-ACT	G	T	E	A	S	A	A	T	A	V	K	I	T	L	L	S	A	L	V
A4; kallistatin	G	T	E	A	A	A	A	T	T	F	A	I	K	F	F	S	A	Q	T
A5; PCI	G	T	R	A	A	A	A	T	G	T	I	F	T	F	R	S	A	R	L
A9; centerin	G	T	E	A	T	A	A	T	T	T	K	F	I	V	R	S	K	D	G
A10; PZI	G	T	E	A	V	A	G	I	L	S	E	I	T	A	Y	S	M	P	P
B1; MNEI	G	T	E	A	A	A	A	T	A	G	I	A	T	F	C	M	L	M	P
B2; PAI-2	G	T	E	A	A	A	G	T	G	G	V	M	T	G	R	T	G	H	G
B3; SCCA1	G	A	E	A	A	A	A	T	A	V	V	G	F	G	S	S	P	A	S
B4; SCCA2	G	V	E	A	A	A	A	T	A	V	V	V	V	E	L	S	S	P	S
B6; PI6	G	T	E	A	A	A	A	T	A	A	I	M	M	M	R	C	A	R	F
B7; megsin	G	T	E	A	T	A	A	T	G	S	N	I	V	E	K	Q	L	P	Q
B8; PI8	G	T	E	A	A	A	A	T	A	V	V	R	N	S	R	C	S	R	M
B9; PI9	G	T	E	A	A	A	A	S	S	C	F	V	V	A	E	C	C	M	E
B10; bomapin	G	T	E	A	A	A	G	S	G	S	E	I	D	I	R	I	R	V	P
B11; epipin	G	T	E	A	A	A	A	T	G	D	S	I	A	V	K	S	L	P	M
B12;	G	T	Q	A	A	A	A	T	G	A	V	V	S	E	R	S	L	R	S
B13; headpin	G	T	E	A	A	A	A	T	G	I	G	F	T	V	T	S	A	P	G
C1; antithrombin	G	S	E	A	A	A	S	T	A	V	V	I	A	G	R	S	L	N	P
D1; heparin cofactor II	G	T	Q	A	T	T	V	T	T	V	G	F	M	P	L	S	T	Q	V
E1; PAI-1	G	T	V	A	S	S	S	T	A	V	I	V	S	A	R	M	A	P	E
E2; protease nexin 1	G	T	K	A	S	A	A	T	T	A	I	L	I	A	R	S	S	P	P
F2; $\alpha_2$antiplasmin[a]	G	V	E	A	A	A	A	T	S		I	A	M	S	R	M	S	L	S
G1; C1-inhibitor[a]	G	V	E	A	A	A	A	S	A		I	S	V	A	R	T	L	L	V
I1; neuroserpin	G	S	E	A	A	A	V	S	G	M	I	A	I	S	R	M	A	V	L
I2; MEPI	G	S	E	A	A	T	S	T	G	I	H	I	P	V	I	M	S	L	A
Noninhibitory																			
A6; CBG	G	V	D	T	A	G	S	T	G	V	T	L	N	L	T	S	K	P	I
A7; TBG	G	T	E	A	A	A	V	P	E	V	E	L	S	D	Q	P	E	N	T
A8; angiotensinogen	E	R	E	P	T	E	S	T	Q	Q	L	N	K	P	E	V	L	E	V
B5; maspin	G	G	D	S	I	E	V	P	G	A	R	I	L	Q	H	K	D	E	L
F1; PEDF	G	A	G	T	T	P	S	P	G	L	Q	P	A	H	L	T	F	P	L
H1; colligin 1	G	N	P	F	D	Q	D	I	Y	G	R	E	E	L	R	S	P	K	L
H2; colligin 2	G	N	P	F	D	Q	D	I	Y	G	R	E	E	L	R	S	P	K	L

[a] For α_2-antiplasmin and C1-inhibitor, the length of the reactive center loop to the P1 residue is only 13 residues. To maintain the alignment of residues within the hinge region as well as at P1−P1′, a gap has been introduced, arbitrarily, at position P6.

Another observation that rate of loop insertion greatly affects the outcome of the reaction comes from a study in which single cysteines were introduced into α_1-PI as sites of covalent fluorophore attachment. It was found that a cysteine at position 298, which is located on the outer face of β-sheet A in strand 5 level with P7 in the covalent complex, could be labeled with fluorophore in the complex without affecting the stability of the complex, but that prelabeling of the native state converted the serpin into a substrate (i.e., decreased k_4 substantially). The fluorophore was presumably acting as a "doorstop" and inhibiting rapid enough passage of the proteinase to achieve an inhibited complex before completion of the substrate reaction.

5.5.2 Importance of Reactive Center Loop Composition

Beyond the general need to have noncharged residues at P14 and nonproline residues at P12 and P10 positions, as described above, there appears to be a more specific sequence restriction within the hinge region of inhibitory serpins, as revealed from sequence comparisons of inhibitory and noninhibitory serpins. In particular, residues P12−P9 show a >50% conservation of alanine at each position (including the P-odd residues), whereas noninhibitory serpins have few or no alanines at these positions, and may incorporate many other types of residue[4,200] (Table 7). This pattern is maintained irrespective of the organism and is also seen, for example, in the set of recently identified *Drosophila* and *C. elegans* serpins (Table 8). P8 also shows a high preference for the small threonine side chain. In the inhibitory serpins the permissible alanine replacements are all relatively small residues, being mostly valine, serine, threonine, or glycine. Although the basis for this pattern of conservation is not known, it seems plausible that it is to facilitate rapid continuation of the insertion process, once the P14 side chain has incorporated into the end of the β-sheet and started the zipper-like process of insertion. The smaller the side chain that needs to be inserted, the smaller the cavity that needs to be created in the interior of the serpin body to accommodate it and so the lower the activation energy for that step is likely to be. Since each extra residue that inserts is also likely to contribute favorably to the thermodynamic stability of the loop-inserted state, loop insertion, once started, is likely to be an increasingly favorable process, both kinetically and thermodynamically, at least for the first five or six residues up to P9, with no build-up of intermediates representing different extents of loop insertion. Such a strong push in the direction of insertion may be necessary to ensure effective irreversibility of the process. Whereas the optimum type of residue for promoting facile insertion is thus

Table 8. *Drosophila melanogaster* and *Caenorhabdis elegans* Reactive Center Loop Sequences for Viable Serpins

serpin	P15	P14	P13	P12	P11	P10	P9	P8	P7	P6	P5	P4	P3	P2	P1	P1′	P2′	P3′	P4′
Drosophila																			
CAB63098	G	S	E	A	A	A	A	T	A	V	V	F	R	Y	K	S	I	R	S
CAB63097	G	A	E	A	A	A	A	T	A	L	L	F	V	R	L	S	V	P	M
CAB63096	G	A	E	A	A	G	A	T	S	V	A	V	T	N	R	A	G	F	S
NM090214	G	T	E	A	A	A	A	T	G	M	I	M	M	T	R	M	M	T	F
NM080218	G	A	E	A	A	A	A	T	A	L	L	F	V	R	L	S	V	P	M
NM080216	G	T	E	A	A	A	A	T	G	M	A	V	R	R	K	R	A	I	M
NM080215	G	S	T	A	A	A	A	T	V	L	F	T	Y	R	S	A	R	P	V
NM080066 (Spn43Aa)	G	C	E	A	A	G	A	S	Y	A	A	G	V	P	M	S	L	P	L
NM080065 (Spn43Ab)	V	T	E	A	G	V	D	Q	P	L	E	T	G	L	L	K	G	L	F
NM080112 (Spn43Ac)	G	T	E	A	S	A	A	S	Y	A	K	F	V	P	L	S	L	P	P
NM080219	G	A	E	A	A	G	A	T	S	V	A	V	T	N	R	A	G	F	S
AAF24518	G	T	E	A	Y	A	A	T	V	V	E	I	E	N	K	F	G	G	S
C. elegans																			
AAB52317	G	T	V	A	A	A	A	T	T	I	S	S	S	V	G	S	V	Q	R
AAB37055	G	T	K	A	A	A	A	T	T	V	S	I	S	L	K	S	A	M	F
AAB52318	G	T	V	A	A	A	A	T	T	I	S	F	S	L	T	S	V	F	I
AAB71720	G	T	T	A	A	A	A	S	A	F	K	V	Q	L	E	M	M	I	M
AAB71272	G	T	R	A	A	A	A	T	E	A	K	I	F	F	T	S	A	S	S
gi3875633	G	T	T	A	A	A	V	T	S	H	N	Y	N	T	L	C	G	A	P

small hydrophobic residues, there are two studies that emphasize the overwhelming importance of P14 in initiating insertion, such that, once started, it requires a major impediment such as proline to slow the process sufficiently to greatly increase the SI. The first was on α_1-PI and involved mutating, inter alia, positions P12, P10, and P8 separately to glutamates.[201] It was found that the SI values were almost unperturbed for positions P10 and P8 and increased only to 4.2 at position P12. A similar "glutamate scan" was carried out on PAI-1 with comparable results,[202] in that, while replacement of the P14 by glutamate converted PAI-1 into a substrate, glutamate substitutions at P12, P10, or P8 still resulted in inhibitors, with a maximum of 30% reduction in apparent second-order rate constant. Given these findings of the absence of a dramatic effect of charged mutations at P-even positions other than P14 in α_1-PI and PAI-1, it is significant that in antithrombin even relatively mild mutations at P12 or P10 have much more serious consequences, with natural mutations to threonine at P12[203] or serine at P10[204] causing decreased inhibition and thrombosis. However, antithrombin is the only serpin clearly shown to have residues P15 and P14 already inserted into β-sheet A in the native state. Thus, there is a very different requirement for antithrombin compared to other serpins. In antithrombin, the reactive center loop composition must *favor* P14 insertion, so that this is the dominant form in the native state, yet this state must be sufficiently stable against further loop insertion that P13 and P12 do not spontaneously insert. This requires that it be hard to insert P12, a situation that would be exacerbated by even mild mutations. The requirement for antithrombin is thus conceptually the same as for other serpins, namely, that it must be hard for the first P-even residue to insert, but once inserted, further insertion becomes increasingly fast and favorable. The difference is that this residue is P14 in most serpins, but P12 in antithrombin.

In the region from P7 to P1, there is no pattern of residue conservation, though the general requirement of hydrophobic residues at the P-even positions is maintained. There are, however, two important considerations here. One is the obvious one of specificity-conferring residues close to P1 being located within this region (discussed below in section 5.6). The second relates to the process of loop insertion toward the ultimate goal of proteinase inhibition. As argued below in section 5.5.4, it may be particularly important for insertion of the second half of the reactive center loop to be highly favorable energetically. A particular residue may be required to give either a large favorable increase in stability of the intermediate as it inserts into β-sheet A or a low activation energy for its insertion. Both of these may be necessary if passage of the proteinase requires temporary displacement of helix F, which normally overlies β-sheet A in the native protein and does so also in the final complex.

5.5.3 Importance of Reactive Center Loop Length

Including the invariant residue that, in loop-inserted forms of cleaved serpins, is at the turning point between strand s3A and the newly inserted s4A (the formerly exposed reactive center loop), the length of the reactive center loop up to the scissile bond is almost always 17 residues (P17−P1). The few documented exceptions are crmA, C1 inhibitor and α_2-antiplasmin, where the length is 16 residues. This invariance is directly related to the inhibition mechanism, which requires that the proteinase be fully translocated to the distal pole of the serpin from the initial interaction site, and that in such a location there be sufficient compression on the proteinase, resulting from it being held by the end of the reactive center loop against the body of the serpin, for the active site to be distorted effectively enough to create the kinetic trap. The length requirement to meet such criteria can thus be considered as the sum of the length needed, upon insertion, to reach the bottom of β-sheet A, plus an amount *less* than that needed to link the carbonyl of the P1 residue to the active site serine 195 of the proteinase, without involving steric clashes between the loops surrounding the proteinase active site and the bottom of the serpin. In this way, accommodation of the proteinase would

necessitate, at a minimum, loop movements to obviate these clashes. In keeping with this, it was noted in construction of the model of the covalent complex derived from FRET data, that just such steric clashes would result from the positioning of the proteinase required by the experimental data and hence would need to be alleviated.[143] Since this length is a function of both serpin and proteinase, one could envision a variation in length requirement for a given serpin, depending on the "thickness" of the loops surrounding the proteinase active site.

An example of critical loop-length dependence is given by α_2-antiplasmin. For arginine-specific proteinases, such as plasmin and trypsin, the specificity-determining P1 arginine is located such that the reactive center loop is only 16 rather than 17 residues long, presumably reflecting a shorter serpin body, such as has been found for another serpin that has a 16-residue requirement, crmA.[93,94] A natural variant, α_2-antiplasmin Enschede, was isolated from an individual with a bleeding disorder.[205−207] In this variant, the insertion of an alanine 8 residues N-terminal of the scissile bond effectively increased the length of the reactive center loop (to the P1 arginine) from 16 to 17 residues. Although the variant retained weak inhibitory properties in a kinetic assay (requiring a large excess of α_2-antiplasmin over plasmin), the complexes were unstable with respect to the SDS denaturation used for SDS−PAGE analysis and, in the kinetic assay, full plasmin activity returned after prolonged incubation. These data suggest both that SI was greatly increased, perhaps as a result of interchange of residues in the P8−P1 region that would be buried in the serpin interior, but more significantly that the stability of the serpin-proteinase complex was greatly reduced, suggesting less effective distortion of the proteinase active site resulting from lengthening of the reactive center loop tether beyond what is required by the thickness of the plasmin loops.

A recent study has more systematically examined the effects of altering loop length within α_1-PI Pittsburgh on the ability to form serpin−proteinase complexes.[208] Shortening of the reactive center loop by up to two residues for inhibition of factor Xa or by one residue for inhibition of thrombin still gave SDS-resistant complexes. Any further shortening resulted in only substrate cleavage. Addition of one residue still allowed complex formation, but with greatly reduced stability, which was reduced even further by addition of a second residue. It should, however, be noted that the proteinase in the complexes formed with a two residue shortened serpin showed no evidence for enhanced proteolytic susceptibility in the complex, suggesting that proteinase compression had not taken place, perhaps as a result of the reactive center loop being *too* short to bring the proteinase to the bottom of the serpin, though perhaps still allowing displacement of the P1 side chain from the S1 specificity pocket and so reducing the rate of deacylation sufficiently to allow trapping of the complex by SDS-denaturation.

Another study has been carried out on crmA, which normally has a P1 aspartate and a reactive center loop length of 16 rather than 17 residues. By mutating P1 to arginine the specificity was changed such that it was recognized by arginine-specific proteinases such as trypsin, thrombin, and factor Xa. Whereas the variant was only a substrate for trypsin, it was able to inhibit and form SDS-stable complexes with both factor Xa and thrombin, though with elevated SI values. Increasing the effective loop length by moving the arginine one residue C-terminal converted the variant into a substrate for all three proteinases (Tesch, Gettins, and Olson, manuscript in preparation).

5.5.4 Kinetics and Thermodynamics of Final Complex Formation

The X-ray structure of the covalent trypsin−α_1-proteinase inhibitor complex provides the clearest evidence that the mechanism of kinetic trapping of the proteinase in the complex is distortion of the active site, as well as of other regions of the proteinase. This distortion, however, must be paid for energetically. The calorimetric study on various cleaved and complexed serpins provides strong evidence that the energy source for this distortion comes from the structural changes within the serpin that result from loop insertion,[110] rather than from a specific set of complementary serpin−proteinase interactions in the final state. This, however, raises an important question concerning the steps involved in achieving the final, kinetically trapped complex. Consider if loop insertion were able to proceed almost, but not quite, to the point of full insertion without any distortion of the proteinase or deviation of the serpin structure from that of the equivalent cleaved, loop-inserted form. The stability of such a state would be much greater than of the desired final complex, the difference in energy being due to that required to distort the proteinase (Figure 9). Thus, any realistic mechanism for introducing such distortion cannot involve smooth insertion of the reactive center loop followed by a spontaneous latching of the proteinase at the bottom and its concomitant distortion, since this would be thermodynamically highly unfavorable. There must somehow be a "coupling" mechanism that makes attainment of the final proteinase-distorted state that has been revealed or implied in the X-ray, NMR, and fluorescence models an energetically favorable one. An obvious candidate for providing the coupling is helix F. This helix and the loop linking it to β-sheet A lie over the lower half of β-sheet A in native, cleaved and complexed structures, obscuring both the site in the native serpin into which the reactive center loop will insert in forming complex and the inserted loop itself in the resulting complex. Thus, even without an attached proteinase, there must be a temporary displacement of this F helix-loop moiety to permit reactive center loop insertion in generating cleaved loop-inserted serpins. With a bulky proteinase attached to the end of the inserting reactive center loop, this situation would be exacerbated and might require much greater displacement of the F helix-loop to permit passage of the proteinase. Only by moving the proteinase "beyond" the end of the sheet (upon achieving complete, or nearly complete insertion) could the F-helix

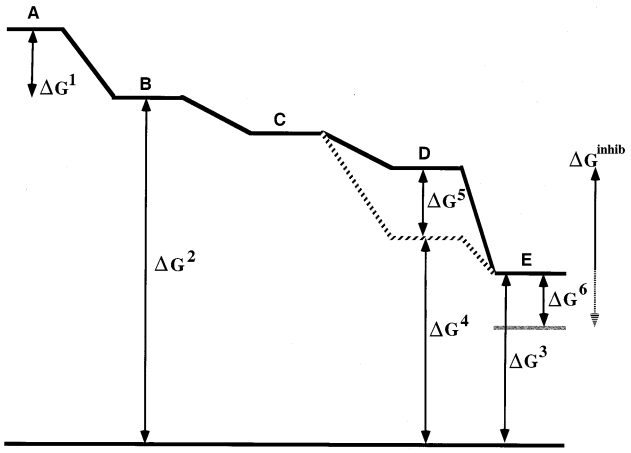

Figure 9. Schematic energy diagram for the coupling mechanism proposed in the text, and shown pictorially in Figure 10, that allows the energy derived from insertion of the reactive center loop to be "stored" in a displaced F helix until needed at the final step to distort the proteinase and ensure that the equilibrium between penultimate and final states greatly favors the latter. The energy derived from loop insertion can be considered to be in two parts, that involving insertion up to about P9 and involving no need to move helix F (ΔG^1) and that for insertion of P8 to P1 (ΔG^2). The energy to displace helix F is ΔG^4, while the energy to pull the P1 side chain from the S1 pocket of the proteinase is ΔG^5, so the energy of loop insertion that is "stored" is ΔG^4 if P1 is not pulled put and $\Delta G^4 + \Delta G^5$ if it is pulled out. ΔG^3 is the total energy required to distort the proteinase, including extraction of the P1 side chain. ΔG^6 is the additional stabilization afforded the final complex if there are specific favorable serpin−proteinase interactions in that state. The energy of stabilizing the final inactive complex (state E) over the penultimate complex (state D), ΔG^{inhib}, is thus $(\Delta G^4 + \Delta G^5) - (\Delta G^3 - \Delta G^6)$, where ΔG^6 is most likely to be dependent on the specificity of the serpin−proteinase pair. Interestingly, ΔG^2, while needing to be large enough to ensure that the reaction proceeds toward the final complex, does not enter into the expression for ΔG^{inhib}.

return to its normal position. Presumably such a displacement of the F-helix away from the favored location would require considerable energy. This displacement, tied to passage of the proteinase could thus provide the coupling to make the final proteinase distortion favorable.

I have recently proposed a mechanism that accommodates the considerations given above.[209] This is depicted in thermodynamic terms in Figure 9 and schematically in Figure 10. The mechanism proposes first that, given the location of the F-helix covering only the lower part of β-sheet A, insertion up to about P9 or P8 can occur with no displacement of helix F, and proceeds smoothly and favorably, with energetics determined solely by the local expansion of β-sheet A (state B). From that point on, the favorable energy of insertion of each additional reactive center loop residue into β-sheet A is partly used to increasingly displace helix F (state C). As long as the net energy change for each insertion is favorable the insertion process will continue in the forward direction (state D). However, much of the insertion energy is not lost but is temporarily stored in the displaced helix F. At the final point of insertion, the ability of helix F to return to its preferred position is coupled to the distortion of the proteinase. As long as the displacement energy is much greater than the distortion energy, the final complex will form spontaneously and create the kinetic trap (state E). An estimate of this has been made for porcine elastase in complex with α_1-PI, where it was found that the difference in stability of free and complexed elastase was about 6 kcal mol^{-1}.[210] Similarly, the energy obtained from insertion of this region of the reactive center loop must be greater than for helix F displacement, so that the process proceeds spontaneously in the desired direction. In this way, helix F behaves like a spring-loaded ratchet.

Of course, this mechanism may be a simplification in proposing that none of the changes seen in the proteinase in the X-ray structure of the covalent complex occurs until the very last reactive center loop residue has been inserted and helix F is able to return to its preferred position. It should also be considered that, as the proteinase moves further down β-sheet A, contacts between the reactive center loop and the proteinase will be progressively lost as a physical requirement imposed by the length of reactive center loop that must be inserted, until finally the P1 side chain must be removed from the S1 pocket. Since all of these interactions, particularly that of P1 with S1, are stabilizing interactions, their loss would require energy input and be part of the total energy cost for later attainment of the final trapped complex. The energy would again come from insertion of each additional loop residue into β-sheet A and would result in a complex of lower stability than if no contacts had been lost. The advantages of such a modification are 2-fold. First, it increases the energy difference for the final step of helix F return and completion of the proteinase distortion, since less energy is required to achieve the final distorted proteinase state. This would shift the equilibrium for this conversion much more toward the final complex. Second, the equilibrium established becomes one between a fully distorted proteinase that has essentially zero catalytic activity and one with only partially restored activity. If all of the changes within the proteinase seen in the X-ray structure (P1 side chain removal and active site residue repositioning as well as domain crushing) only come at the very end, the equilibrium would be between a zero activity state and one with full proteinase activity. With the resulting smaller energy difference between D and E states if there were no proteinase distortion in the D state, there would also be a higher fraction in the active D state. Both factors would reinforce to compromise the effectiveness of the inhibition process, whereas the reverse would be true for partial loss of catalytic efficiency prior to the final compression.

Evidence in support of this mechanism comes from the recent X-ray structure of the cleaved form of an ovalbumin P14 hinge region variant.[95] The mutation of P14 arginine to threonine allows full insertion of the reactive center loop when cleaved at P1−P1′. The variant is still not, however, capable of inhibiting proteinase. Comparison of the native and loop-inserted cleaved structures indicated that helix F and the strand that connects it to β-sheet A have particularly stable interactions in the native state and

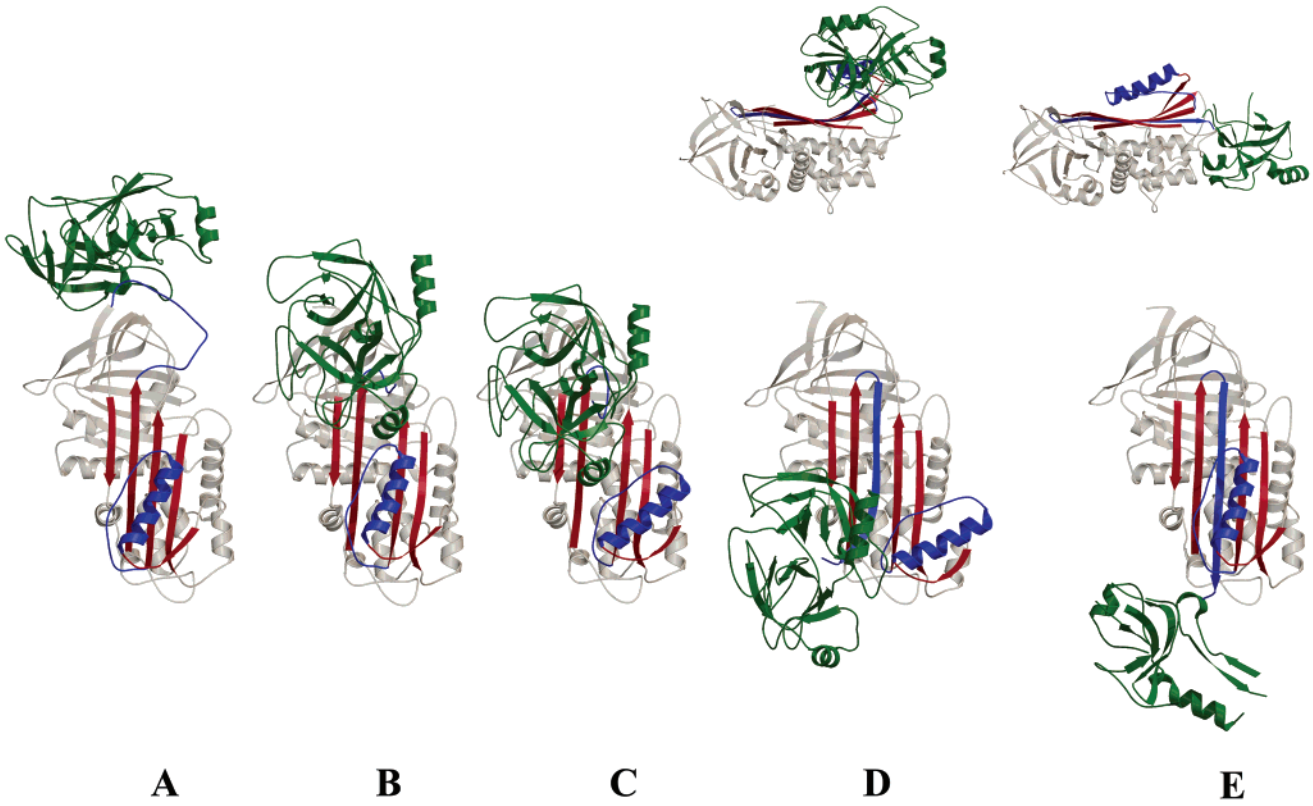

Figure 10. Schematic of the structures involved in the energy diagram of Figure 9 and described in the text. Panel A depicts the situation at the point of cleavage of the RCL and prior to loop insertion. Panel B depicts the complex after the first few RCL residues have inserted and just as the proteinase reaches helix F. Panel C shows the partial displacement of helix F (in blue) needed to permit continued progress of the proteinase (in green) toward the bottom of the serpin. Panel D shows full displacement of helix F with removal of the P1 and other side chains of the reactive center loop from contact with subsites on the proteinase. Above panel D is a side-ways view of structure D to show that the proteinase is still "above" the plane of β-sheet A and otherwise not distorted. This represents the penultimate state in which the proteinase is no longer fully active as a result of the reorientation of the acyl intermediate within the active site as a consequence of removal of P1 from its specificity pocket. Panel E shows the final complex in which the F helix has returned and the proteinase has been distorted through compression against the bottom loops of the serpin, and consequently been fully inactivated. The energy needed to bring about this compression in an overall favorable process comes from return of the F-helix to a location where its interactions are optimal. Above panel E is a side-ways view of the structure to show the movement of the proteinase below the plane of β-sheet A and against the bottom of the serpin. Reproduced with permission from Gettins, P. G. W. *FEBS Lett.* **2002**, *523*, 2. Copyright 2002 Elsevier Science N. V.

would likely be difficult to move. This mechanism also provides a nice explanation for an antibody mapping study that at first appearance is in contradiction to the X-ray, NMR, and fluorescence-based structures of the serpin−proteinase complex. In this study, a monoclonal anti-antithrombin antibody was characterized that bound tightly, and with similar affinity, to two different antithrombin-proteinase complexes, but not to native, cleaved or latent antithrombin.[211] The epitope was found to reside entirely in s4A, at a position that is covered by helix F and the loop connecting it to s3A in all of the monomeric structures, and by analogy with the α_1-PI-trypsin covalent complex structure, in the covalent complexes as well. However, my hypothesis predicts that the "final structure" is an equilibrium between states D and E, with the equilibrium favoring E, but not necessarily by more than a few kcal mol^{-1} (see below). A high affinity antibody could thus still bind tightly to the covalent complex by binding to the D state, with use of some of the binding energy to push the equilibrium back toward state D. Consideration of Figure 9 shows that this would require much more energy in any of the uncomplexed states of antithrombin, since the displacement of helix F is not partially compensated by relief of proteinase distortion. Antibody binding would thus be many orders of magnitude weaker to these species. Additional evidence in support of the hypothesis is considered in the next section.

5.5.5 Stability of the Covalent Complex

In considering the "stability" of the serpin−proteinase covalent complex, it must always be remembered that one is considering the kinetic stability of an acyl enzyme intermediate rather than the thermodynamic stability of an end product. The kinetic stability is with respect to the final products of the reaction, namely, cleaved loop-inserted serpin and regenerated free proteinase. Thus, aspects of stability that are considered here are those that determine how quickly the acyl intermediate is hydrolyzed to the final end products.

An important conclusion from the above model of how translocation of the proteinase and insertion of the reactive center loop into β-sheet A might be favorably coupled to proteinase distortion is that the final state, in which the proteinase is grossly dis-

torted, is in equilibrium with the immediately preceding state, in which the proteinase has lost important interactions with the P1 residue. These different states for the proteinase would be expected to have very different catalytic efficiencies, and so would determine the overall stability of the covalent complex on the basis of what the fractional occupancy and activity of each state is.

Very nice support for the existence of such an equilibrium and for the markedly different properties of the contributing states has come from a recent study on the pH-dependence of serpin−proteinase dissociation for several serpin−trypsin pairs[151] and a study of α_1-antichymotrypsin−HNE complexes of varying stability.[212] The first study showed that the dominant form of the serpin−proteinase complex is deacylated in a manner that is independent of His-57 of the catalytic triad of the proteinase, and depends only on noncatalyzed, hydroxide-mediated hydrolysis. The acyl intermediate is partially protected from solvent, so that added hydroxylamine, a larger nucleophile than hydroxide, has only restricted access. The properties of this state correspond well with those expected for the grossly distorted state seen in the X-ray crystal structure in which His-57 has moved 6 Å away from the acyl linkage. However, in the presence of Ca^{2+}, which binds to one of the active site loops of trypsin that are disordered in the X-ray structure of the covalent complex, the rate of complex dissociation was greatly increased (up to 80-fold) and the pH-dependence of the dissociation reverted to that characteristic of catalytic involvement of His-57. These data could be fitted to a two-state equilibrium between a dominantly populated inactive state and a minor low activity state, with the effect of Ca^{2+} being to shift the equilibrium toward the low activity state. For the antithrombin−trypsin pair, it was further possible to calculate that the catalytically active state is ~0.2% populated at pH7.2, but with a specific activity (2×10^{-4} s^{-1}) that, although orders of magnitude faster than the hydroxide-mediated hydrolysis (5×10^{-7} s^{-1}) is still several orders of magnitude less than would occur for this step in a normal substrate cleavage deacylation (>10−100 s^{-1}, based on measured rates of acylation and the assumption that deacylation is not the rate limiting step for the whole reaction). In addition, the enhanced proteolytic susceptibility of the proteinase in the fully distorted state of the complex is absent from the partially active state, suggesting a much smaller degree of structural perturbation. The properties of the active state of this study thus correspond exactly with those proposed by the present model, namely, that some of the changes within the proteinase that result in loss of catalytic efficiency have already occurred in the steps immediately preceding the final step. In addition, the ΔG between the two end states is modest (~3.7 kcal mol^{-1} calculated from their relative occupancies). In terms of gross physical changes in conformation, the partially active state would correspond to one in which the F-helix is still displaced by the proteinase, the reactive center loop is fully inserted and therefore taught with respect to the active site serine, the P1 side chain has been extracted from the S1 pocket, but the oxyanion hole has not yet completely collapsed as a result of proteinase compression.

The second study used the pH-dependence and solvent isotope effects on the breakdown of α_1-antichymotrypsin−HNE pairs of varying stability to probe the mechanisms involved in breakdown. Here, by use of different RCL variant α_1-antichymotrypsins, complexes were formed that were either quite unstable, with half-lives from minutes to hours, or had normal long term stability. The behavior of the dominant species was examined in each case. In the case of the stable complex, this represented the fully inactivated form (state E in Figures 9 and 10), which showed deacylation that was independent of the pK_a of the catalytic histidine. For the less stable complexes, the equilibrium between penultimate and final states had presumably been shifted much more toward the penultimate state (state D in Figures 9 and 10). As with the antithrombin−trypsin and α_1-PI−trypsin pairs, there was a pH dependence of deacylation that indicated involvement of the catalytic apparatus of the proteinase, albeit at a reduced rate, consistent with the penultimate state being functional as an enzyme, but compromised in its effectiveness.

In the context of the above two-state, two-activity model of serpin−proteinase complexes, the question of complex stability is reduced to one of addressing what determines the population of the partially active state and what might influence the activity of this state. According to the energy diagram of Figure 9, the ΔG between the two conformations (ΔG^{inhib}) is determined by several energies. One is the energy required to displace the F-helix (ΔG^4). A second is the energy to fully distort the proteinase (ΔG^3). The third is the energy already expended prior to the return of the F-helix to partially compromise the proteinase through P1 side chain extraction (ΔG^5), and the fourth is any additional stabilization or destabilization caused by bringing together the proteinase with the bottom of the serpin in the final state (ΔG^6). As suggested by the calorimetric study on the enthalpy changes resulting from serpin−proteinase complex formation versus simple loop insertion, the value of ΔG^3 may be rather large. Since ΔG^4 must be larger than ΔG^3 for the mechanism to move to completion, the net energy difference may be the difference of large numbers and therefore difficult to predict with accuracy.

For a given serpin and a range of proteinases, the main variables in determining relative stabilities of the different complexes are likely to be the energies required to distort the different proteinases and any energy that might result from specific serpin−proteinase interactions in the final state (ΔG^6). Other things being equal, a serpin is more likely to have evolved to have a favorable ΔG^6 contribution with a specific target proteinase than with an unnatural one. In keeping with this, antithrombin−thrombin and antithrombin−factor Xa complexes were found to be much more stable than antithrombin−trypsin complexes.[151] Similarly, for a serpin reacting with several unnatural proteinases, the order of complex

stability is likely to be inversely related to the stability of the proteinase. Unfortunately, there are no reliable experimental data at present to test this prediction.

5.5.6 Proteolysis of the Proteinase in the Covalent Complex

It has long been known that the proteinase moiety in covalent serpin−serine proteinase complexes has much higher susceptibility to proteolysis than does the free proteinase.[213,214] This is readily understandable in terms of the X-ray crystal structure of the α_1-PI−trypsin complex, which shows sufficiently high disorder for ~40% of the trypsin moiety (the region containing the sites of enhanced proteolytic susceptibility) that it suggested a loosening of the structure, perhaps akin to formation of a molten globule like structure. On SDS−PAGE, this is seen as bands with mobilities intermediate between those of the first formed covalent complex and of the cleaved serpin. Whereas this has been correctly used as a demonstration of an altered state for the proteinase in the complex,[181,183] it has also been recognized that it may have physiological significance, since such proteolyzed complexes do not release active proteinase upon deacylation[139,182,215] and so may serve as a means of permanently inactivating the proteinase, even if it were to dissociate from the serpin−proteinase complex. Under in vitro conditions, the proteinase responsible for proteolysis of the complex is typically the same proteinase used for complex formation and present in excess. However, in vivo the proteinase may be any present, and need not be restricted to serine proteinases. Neutrophil elastase, present in high concentration at sites of inflammation, might be particularly important for such inactivation.[182]

The apparent susceptibility of all serine proteinase−serpin complexes to such enhanced proteolysis in the proteinase moiety suggests not only a common type of structural transformation for different proteinase in forming kinetically trapped serpin complexes, but also a common dramatic change in biochemical properties. Since serine proteinases of the trypsin family are all derived from low or zero activity zymogen forms, one can envision a much simpler type of structure for the proteinase in the covalent serpin complex than that found, in which the proteinase has reverted to a zymogen-like state. This had indeed been proposed prior to elucidation of the X-ray structure.[109] That this simple solution is consistently *not* used supports the idea that the enhanced proteolytic susceptibility is a necessary feature that has been deliberately evolved for full in vivo functioning, and that is in addition to the rapid kinetic inactivation of the proteinase in the complex.

In this regard, the behavior of the only serpin-cysteine proteinase reaction that has been well characterized (that between papain and antithrombin−see section 5.8) is even more intriguing. Analysis of the reaction showed that, while an analogous complex to that found for serine proteinases appeared to form, with generation of a thioacyl intermediate, loop insertion within the antithrombin molecule and enhanced proteolytic susceptibility of the complexed proteinase, the increase in proteolytic susceptibility was so large that the papain was completely degraded to small peptides.[216] Rather than being an artifact of the use of non-cognate proteinase and serpin, it may reflect the normal consequences for proteinases with the papain fold (e.g., the intracellular cathepsins K, L, and S). Such complete degradation might be desirable for intracellular cysteine proteinases, since there are no intracellular clearance receptors equivalent to LDLR family members to clear such serpin:proteinase complexes from the cytoplasm, other than targeting to the proteosome.

5.6 Determinants of Proteinase Specificity

5.6.1 Concepts of Inhibitory Effectiveness and Specificity of Serpins

When dealing with nonserpin serine proteinase inhibitors, the effectiveness of an inhibitor is determined by the affinity with which it binds to the target proteinase. Selectivity for one proteinase over another is determined by the relative affinities of binding. For such inhibitors it makes sense to describe the effectiveness in terms of a K_I or an IC_{50}. For serpins, it is important to make a distinction between "effectiveness" as an inhibitor and its specificity for a given proteinase. Once a serpin has formed a covalent, kinetically trapped complex with a target proteinase, its effectiveness as an inhibitor is essentially 100%, since the only practical way that free proteinase can be regenerated is by progression further to completion of the hydrolysis reaction, with very slow rate constant k_5, rather than by reversal of the pathway up to that point. Since formation of the trapped complex occurs as the product of one branch of a two branch pathway, a more useful concept of effectiveness as an inhibitor is the value of SI. Thus, a serpin with SI of 1.00 would be the most effective inhibitor, whereas one with an SI of 100 would behave essentially as a substrate for the proteinase. SI is not, however, concentration dependent, and so it is meaningless to describe the effectiveness of a serpin, as defined here, in terms of a K_I. Conversely, the apparent second-order rate constant for reaction of a serpin with a proteinase to form an acyl intermediate complex (k_2/K_M), while reflecting the specificity of the reaction, gives no information on the effectiveness as an inhibitor (SI). Two proteinases may react with the same serpin with second-order rate constants that differ by several orders of magnitude and yet both might have an SI ~ 1. As long as both reactions were allowed to go to completion, each proteinase would be inhibited with the same effectiveness, yet at intermediate time points one proteinase would be inhibited to a much greater extent than the other, in a manner dependent on the second-order rate constants and the concentrations of the reacting components.

In the older literature and, regrettably even now, too little regard is sometimes given to these different aspects of serpin inhibition. Thus, a serpin mutated at a given position into a series of variants may be incubated for a fixed time with a proteinase and the amount of residual proteolytic activity remaining for each variant be reported either as a measure of the

Table 9. Influence of RCL Residues on Proteinase Specificity of Serpins

serpin	position	residue	proteinase specificity	other comments	ref
α_1-PI	P1	Met	neutrophil elastase porcine elastase	wild-type	26
	P1	Arg	thrombin, plasmin	Pittsburgh variant	218
	P1	Phe	cathepsin G		219
	P1	Leu	neutrophil elastase, cathepsin G		219
	P3, P1	Ala, Val	porcine elastase		219
	P2−P1−P1′	Pro-Arg-Thr	thrombin	attempt to create a thrombin-specific inhibitor	483
α_1-antichymotrypsin	P1	Leu	cathepsin G	wild type	414
	P1	Trp	chymase		414
	P1	Arg	trypsin		484
	P1	Met	elastase		485
kallistatin	P1	Phe	tissue kallikrein, chymotrypsin	wild type	221
	P1	Arg	tissue kallikrein, trypsin		221
PCI	P1	Arg	protein C, thrombin	wild type	486
	P1	Met	chymotrypsin		487
	P2	Pro	thrombin	reduced inhibition of APC	488
SCCA1	P1	Gly	cathepsin S	one residue shorter RCL	303
	P1	Arg	trypsin		303
SCCA2	P1	Leu	cathepsin G		310
	P1	Ser	cathepsin S		310
antithrombin	P1	Arg	thrombin, factor Xa	wild type	489
	P1	Trp	chymotrypsin		230
	P2	Pro	thrombin	thrombin:factor Xa ratio increased	490
	P1′	Leu	thrombin, factor Xa	Denver mutation. Reduced rates of conversion of noncovalent complex to acyl intermediate	327
heparin cofactor II	P1	Leu	thrombin, chymotrypsin	wild type	491
	P1	Arg	thrombin, factor Xa		492
PAI-1	P1	Arg	uPA, tPA	wild type	493
	P1−P1′	Lys-Ala	uPA≫tPA		494
	P1−P1′	Lys-Trp	tPA>uPA		494
α_2-antiplasmin	P1	Arg	plasmin	wild type, 1 residue shorter RCL	27
	P1	deleted	elastase	recognition of Met at P1′ as new P1	495
C1 inhibitor	P1	Arg	C1s, kallikrein	wild-type	496
	P1	Cys	not inhibitory	Da mutant, causes angioedema	497

effectiveness of the variant serpin as an inhibitor, or of the selectivity conferred by that mutation. However, the remaining proteolytic activity is a function both of the rate of reaction selectivity and of the SI (effectiveness as an inhibitor). Without explicitly determining each component, the reported residual activities are neither a direct measure of effectiveness nor of specificity.

Factors that influence effectiveness (SI) are dealt with in other sections. Here I consider specificity, i.e., those factors that influence steps leading up to formation of the acyl enzyme intermediate and that are reflected in the term k_2/K_M. In this regard, specificity of serpins is determined by the same types of consideration as for specificity of substrates for proteinases, i.e., k_{cat}/K_M. As is illustrated below, however, the specificity of reaction of the same sequence as a free peptide and as a stretch of the reactive center loop of a serpin may be quite different, resulting from additional contributions to K_M in the serpin.

5.6.2 Residues within the Reactive Center Loop

Both the NMR structural study and the X-ray crystal structure of noncovalent Michaelis complexes of serpin and proteinase discussed above agree that the initial docking does not involve any insertion of the reactive center loop into β-sheet A or any induced fit in the proteinase. Therefore, whether there is only a single pattern of recognition between all serpins and their target proteinases or many, it seems likely that primary contributions to specificity will be determined by residues within the reactive center loop that interact with subsites on the proteinase, in and around the active site. The most direct measure of the influence of a particular residue on K_M is the dissociation constant of the noncovalent complexes it forms with the inactivated proteinase. Much more common, however, is to determine second-order rate constants for the inhibition reaction and treat changes in this value as a measure of change in K_M, i.e., assuming that k_2 has been unchanged by the mutation.

Many studies have been carried out on the effects of mutations at the P1 position or adjacent to this to understand the nature of specificity in a serpin−proteinase reaction (Table 9). Such mutations include both naturally occurring ones and ones introduced by site-directed mutagenesis. An example of the former is the Pittsburgh variant of α_1-proteinase inhibitor in which the P1 residue had mutated from methionine to arginine. This led to a change in both absolute and relative rates of reaction toward poten-

tial target proteinases Whereas the wild-type methionine reacts with its normal target, neutrophil elastase, with second-order rate constant of 6.5×10^7 M^{-1} s^{-1},[26] the arginine variant does so with rate constant 2.2×10^3 M^{-1} s^{-1}.[217] More significant is that the rate of reaction with thrombin increases from a mere 48 M^{-1} s^{-1} to 3.1×10^5 M^{-1} s^{-1},[217] making it a much faster inhibitor of thrombin than antithrombin (in the absence of heparin). Not surprisingly, the mutation was identified as a result of major hemorrhage in the affected individual, which eventually led to death.[218]

An early example of site-directed mutagenesis applied to studying the effects of P1 mutations was on recombinant α_1-proteinase inhibitor, in which the P1 methionine was changed to valine, isoleucine, alanine, leucine, phenylalanine, or arginine.[219] Here, changes in P1 alone could significantly alter the absolute and relative rates of reaction with different target proteinases. Replacement of P1 Met with Leu gave even faster reaction with neutrophil elastase and enhanced inhibition of cathepsin G. Replacement of P1 with Phe gave a specific cathepsin G inhibitor, while replacement with Ala, Ile, or Val resulted in efficient neutrophil elastase inhibitors, but not of cathepsin G.

A demonstration that additional residues can be important in determining rate of reaction and proteinase specificity was a study that started with a P1 arginine variant of α_1-antichymotrypsin (which made it an effective inhibitor of thrombin and trypsin) and subsequently introduced different residues at the P2 position.[220] Seven different residues were introduced at the P2 position and the rate constants for reaction with various arginine-specific proteinases determined. In the case of thrombin, the nature of P2 altered the measured $k_{apparent}$ by up to 3 orders of magnitude. A problem, however, with this and many other early studies is that a correction for SI was not made in reporting second-order rate constants.

In the case of the P3 position, a study has been carried out that includes effects of P3 variation in the reaction of the serpin kallistatin with the potential target proteinase tissue kallikrein.[221] Here, however, only a semiquantitative measure of rate of reaction was obtained from densitometric scans of SDS−polyacrylamide gels of complex formation after a fixed period of incubation. While lysine (wild-type residue) and arginine appeared to give the highest degree of completion of reaction, leucine, alanine, glutamine, glycine, histidine, and methionine were almost as effective, suggesting that P3 plays only a minor role in influencing specificity of kallistatin. The role of residues on the C-terminal side of the reactive bond has also been examined. In a study on the substrate preference of thrombin for different residues at P2′ and P3′, using fluorogenic decapeptide substrates,[222] it was found that phenylalanine at P2′ in place of aspartate enhanced k_{cat}/K_M for the reaction by 3 orders of magnitude. P3′ was less important, but thrombin still showed a preference for positively charged side chains at this position and disfavored acidic residues.

Two major conclusions from many mutagenesis studies aimed at determining which residues in the reactive center loop of a serpin contribute to proteinase specificity are (i) P1 is usually the most important residue, followed by P2 and then P3 and (ii) by appropriate change of these residues a serpin can be designed to inhibit (by the same conformational change mechanism) a quite different proteinase. This is potentially of great use in creating "designer" serpins, in that, as long as the length of the reactive center loop between the hinge point and the reactive bond is appropriate (see 5.5.3), incorporation of the correct residues into the reactive center loop might enable an optimal proteinase specificity to be introduced. Of course, the situation may be separately complicated by the effect that multiple contiguous mutations have on loop insertion and hence on SI, and on complex stability. An example of this is the attempt to transform α_1-PI into an optimal inhibitor of furin.[223] Consideration of the requirements for substrate cleavage by furin showed that arginine is required at P1, and that at least two of the three positions P2, P4, and P6 should be arginine or lysine. While introduction of arginine at P2 (in addition to arginine at P1) gave a furin inhibitor that was very much more effective than the P1 arginine variant alone,[224] further modification of the reactive center loop sequence to incorporate additional arginines at P4 and P6 gave little further rate enhancement, and instead served to exacerbate an already poor SI, increasing it from 3.6 in the case of the P2−P1 arginine double mutant to 24 with the RERIRR P6−P1 sequence.[223]

5.6.3 Exosite Interactions

In addition to the usually dominant importance of P1 and other nearby residues in the reactive center loop in determining proteinase specificity, there are also at least two examples of exosites playing an important additional role. The first is in the α_1-PI/human neutrophil elastase (HNE) reaction. This has a second-order rate constant of $\sim 6 \times 10^7$ M^{-1} s^{-1} and an SI of $\sim$1.[26] An attempt to change the specificity of α_1-antichymotrypsin to that of α_1-PI by progressive replacement of increasing stretches of its reactive center loop with those of α_1-PI (centered on the P1−P1′ pair) gave very poor results.[225] Whereas change of the P1 from leucine to methionine did convert it from a substrate of HNE to an inhibitor with SI of 5, the second order rate constant was only 4×10^4 M^{-1} s^{-1}. Change of the entire region from P3 to P3′ increased the rate constant only another 2.5-fold, and reduced the SI in proportion. Thus, the best chimeric inhibitor was still about 60-fold slower than α_1-PI at inhibiting HNE, strongly suggesting the importance of either serpin body−proteinase interactions or serpin body-reactive center loop interactions in determining specificity. This conclusion is reinforced by a much earlier study that examined the reactivity of several series of peptides related to the reactive center loop sequence of α_1-PI. It was found that, even with the identical P4−P4′ sequence that is present in the serpin, the rate of reaction (k_{cat}/K_M) was only 1×10^4 M^{-1} s^{-1}, i.e., 3 orders of magnitude less than in the intact serpin.[226] In addition, the values of K_M

for the peptides, while showing some variation as a function of peptide length and composition, were all weak and in the mM range, suggesting that, to obtain the rates of reaction found in α_1-PI, additional exosite interaction would be necessary. At present no further details are available of where this exosite is or what the nature of the interactions might be.

The second example of exosite involvement is in the reaction of factors IXa and Xa with heparin-activated antithrombin. As described in more detail in section 5.7, heparin species containing a specific pentasaccharide bind to antithrombin and induce significant conformational changes within the molecule. The conformationally altered antithrombin reacts with both factors IXa and Xa at rates about 300-fold faster than does native antithrombin uncomplexed with heparin.[227,228] Until recently, this rate acceleration was thought to be due to conformational changes within the reactive center loop altering the flexibility and complementarity of the reactive bond sequences for the proteinase. From recent extensive mutagenesis studies, however, it is now clear that this is not the case and that instead there must be an exosite in heparin-complexed antithrombin that is not present in native antithrombin and that provides additional binding interaction in the initial noncovalent complex.[229,230] Thus, studies involving mutation of the P1 residue gave, not unexpectedly, antithrombins with different proteinase specificities (the change to tryptophan, leucine, or methionine greatly increased the rate of inhibition of chymotrypsin at the expense of reaction with factor Xa, trypsin, or thrombin), but ones in which the additional enhancement of inhibition of factor Xa that results from the heparin-induced conformational change was not abrogated.[229] More extensive mutations were made within the reactive center loop in an attempt to change the specificity of antithrombin from factor Xa to thrombin. Although this was successful, it was surprisingly found that the heparin-induced conformational change was still minimally important for thrombin inhibition in all of the variants, but was of almost unchanged high importance for factor Xa inhibition. It now appears that antithrombin also shows a similar heparin-induced exosite-dependent acceleration of inhibition of factor IXa.[228] Thus, of four proteinases that have been extensively studied so far in their reactions with heparin-antithrombin, trypsin, and thrombin show no involvement of exosite interactions with the serpin body, whereas factors IXa and Xa do. It may turn out that other coagulation proteinases also use exosites in their interaction with heparin−antithrombin.

5.7 Regulation of Activity

5.7.1 Advantages of the Serpin Mechanism

An important question concerning the widespread use of serpins, rather than rigid lock-and-key inhibitors, by many species in many situations is why they are used when they are so prone to malfunction as a result of even single residue mutations in many parts of the protein, and when they involve single turnover of the serpin as an inhibitor. Although the examples are still limited, it is very likely that the answer lies in the possibilities for regulation of activity and outcome that can occur with serpins, but cannot occur with reversible inhibitors. Thus, consideration of the branched pathway mechanism shows that it might be possible to modulate both the rate of reaction of the serpin with the proteinase (by affecting k_2/K_S) and the outcome of the reaction (by modulating k_4 relative to k_3). In addition, since the reactive center loop may be cleaved nonproductively at sites other than P1−P1′ by nontarget proteinases, there is the possibility of controlling the level of functional serpin by proteolytic inactivation. In this section, known examples of such regulation are considered in detail, the activation of certain serpins by glycosaminoglycans, including heparin, the regulation of PAI-1 activity by binding to vitronectin, the cofactor role of protein Z in the inhibition of factor Xa by protein Z-dependent proteinase inhibitor, and the proteolytic inactivation of serpins by nontarget proteinases.

5.7.2 Heparin and Other Glycosaminoglycan Activation Mechanisms

Several serpins found in vertebrate blood plasma are activated as inhibitors of target proteinases by binding to heparin or other linear negatively charged glycosaminoglycans. These include antithrombin, heparin cofactor II, PAI-1, protein C inhibitor, and protease nexin 1, a group of serpins involved in regulation of proteinases of the blood coagulation and fibrinolysis systems. The resulting enhancements in the rates of proteinase inhibition can be up to several 1000-fold. Such rate enhancements, taken together with the abundance of these glycosaminoglycans at sites of serpin action, suggest an important physiological role for such activation in site-specifically regulating the inhibitory action of these serpins.

There are two principal types of activation mechanism used by glycosaminoglycans: (i) a bridging mechanism in which the linear glycosaminoglycan serves to simultaneously bind both serpin and proteinase and thus bring them together in an appropriate orientation for productive interaction of the serpin reactive center loop with the proteinase active site, and (ii) a conformational change-based mechanism in which glycosaminoglycan binding to the serpin serves to alter the serpin's conformation to one in which it is more reactive toward the proteinase. In the case of PAI-1, protease nexin 1 and protein C inhibitor, only the bridging mechanism is used, while with antithrombin and heparin cofactor II both bridging and conformational change can be used. Table 10 gives a summary of the contributions of the two types of mechanism to inhibition of selected proteinases by these five serpins.

Bridging by Glycosaminoglycans. A requirement for this mechanism is that there be GAG binding sites on both serpin and proteinase component, with appropriate relative locations and orientations to allow optimal alignment of the two proteins for reaction with one another when bound to the GAG. A common feature for four of the five serpins considered here (antithrombin, heparin cofactor II,

Table 10. Bridging and Conformational Change Contributions of GAGs to Rate Acceleration of Selected Serpin−Proteinase Reactions

serpin	proteinase	GAG	bridging	conformational	max rate (M^{-1} s^{-1})	ref
antithrombin	thrombin	heparin	~2400-fold	1.7-fold	3.7×10^7	242
antithrombin	factor Xa	heparin	70-fold	300-fold[a]	4.4×10^7	227, 243
antithrombin	factor IXa	heparin	600-fold	700-fold[a]	2×10^7	228
antithrombin	trypsin	heparin	2.2-fold	3.2-fold	1.4×10^6	229
HCII	thrombin	heparin	7-fold	~2400-fold	1.1×10^7	275, 441
HCII	thrombin	dermatan sulfate	no effect	~17000-fold	1.1×10^7	275, 441
protease nexin 1	acrosin	heparin	500-fold[b]	N. D.	2.4×10^6	154
protease nexin 1	activated protein C	heparin	44-fold[b]	N. D.	2.3×10^5	498
protease nexin 1	thrombin	heparin	900-fold[b]	N. D.	1.2×10^9	446
protease nexin 1	Factor Xa	heparin	70-fold[b]	N. D.	3.5×10^5	499
protease nexin 1	Factor XIa	heparin	215-fold[b]	N. D.	1.7×10^6	347
PAI-1	thrombin	heparin	200-fold	no effect	1×10^5	443, 500
PCI	acrosin	heparin	230-fold	possible 2-fold negative effect on acrosin	5.6×10^7	420
PCI	thrombin	heparin	10-fold	N. D.	1.9×10^5	417

[a] In the presence of Ca^{2+}, which binds to the Gla domain of Factors IXa and Xa. [b] Since the bridging and conformational contributions were not separately determined, the total enhancement is accorded to bridging.

PAI-1, and protease nexin 1) is the involvement of basic residues within the D helix in the heparin binding site, suggesting a common evolutionary origin. Antithrombin additionally uses basic residues in the adjacent A helix and in a region immediately N-terminal to the D-helix that forms a new short helix upon heparin binding (the P-helix).[231] The resulting heparin binding site in antithrombin not only has the highest heparin affinity of these serpins (10−20 nM),[232−234] but also is the only one with specificity for a particular heparin sequence (a specific pentasaccharide).[235−239] Protein C inhibitor is the outlier, and instead has been shown by modeling and mutagenesis studies to use basic residues in helices A and H.[240] Of the target proteinases, factor IXa followed by thrombin have the tightest heparin binding sites,[241] and show the largest contributions from bridging to the acceleration of proteinase inhibition,[242−244] though conservation of at least some of the basic residues that form the heparin-binding exosite of thrombin in other proteinases of the same chymotrypsin family results in some weak heparin affinity for other proteinases and correspondingly some bridging contribution to GAG-mediated activation (Table 10). An illustration of the way in which such bridging results in bringing together serpin and proteinase components is given by a model of a ternary heparin−antithrombin−thrombin complex, based on structures of each component and knowledge of the location of the heparin binding sites on each (Figure 11). In this model, it was apparent that the central portion of the heparin chain is not in contact with either heparin binding site and so might be replaceable by a nonheparin linker. A test of this, using a synthetic oligosaccharide consisting of the specific high affinity antithrombin pentasaccharide linked via a nonsulfated spacer to a nonspecific sulfated pentasaccharide, elegantly demonstrated the validity of this hypothesis.[245]

A characteristic of such a bridging mechanism of activation is that there is a bell-shaped dependence of the rate acceleration on the heparin concentration. This results from binding of both serpin and proteinase to the same GAG chain at low GAG concentrations, but binding of each protein to a separate chain

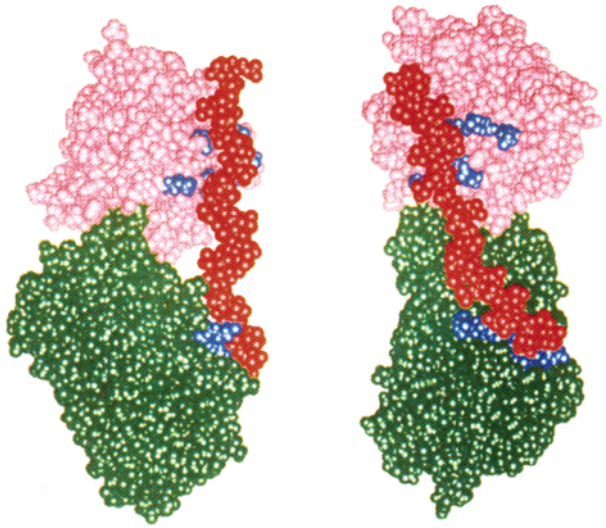

Figure 11. Model of the antithrombin−thrombin−heparin ternary complex, showing orthogonal views. Antithrombin is in green, thrombin is in magenta and heparin is in red. The heparin binding residues on antithrombin and thrombin are shown in blue. The model was constructed by docking the X-ray structures of antithrombin and thrombin with an NMR-derived structure of heparin. Reproduced from Grootenhuis, P. D. J.; Westerduin, P.; Meuleman, D.; Petitou, M.; van Boeckel, C. A. A. *Nat. Struct. Biol.* **1995**, *2*, 736−739. Copyright (1995) Macmillan Magazines, with permission.

at very high GAG concentrations. This has been demonstrated for reactions of thrombin with antithrombin,[246] with PAI-1,[247] with protease nexin 1,[248] and with heparin cofactor II.[249] An additional expectation of such a bridging mechanism is a minimum length-dependence for the activating GAG, to accommodate both protein components on the same polysaccharide chain. For antithrombin−thrombin reactions this has been shown to be heparin ≥ 18 saccharides in length,[227,250,251] for PAI-1-thrombin heparin >14 saccharides in length,[247] and for heparin cofactor II-thrombin heparin >24 saccharides in length.[252]

Conformational Change Induced by GAGs. For two of the GAG-activatable serpins, antithrombin and heparin cofactor II, binding of GAGs additionally induces a conformational change within the serpin that can also contribute to the increase in rate of

proteinase inhibition, though this contribution is also proteinase-dependent. Since the way in which each of the two serpins behaves is very different, each is considered separately.

a. Antithrombin. The X-ray structures of antithrombin determined in 1994 showed that, in contrast to other serpins, the residues at the hinge of the reactive center loop, P15 and P14, are already inserted into β-sheet A in the native, low activity state.[88,89] It was proposed that heparin binding causes extension of helix D, which is connected to β-sheet A by a short linker, and that such helix extension causes contraction of β-sheet A and expulsion of the hinge residues P15 and P14.[253] The altered flexibility and conformation of the reactive center loop were then proposed to represent an activated state for optimal interaction with factor Xa, thereby explaining the conformation-dependent contribution of heparin binding to acceleration of the rate of inhibition of this proteinase.[227] An X-ray crystal structure of the complex with the specific high affinity heparin pentasaccharide, provided some, but not all, of the answers to whether this model of heparin activation is correct.[231] The structure showed that heparin binds to basic residues in and around helix D, induces residues 112−119 to adopt a helical conformation so that lysine 114 can also interact with the pentasaccharide, and causes the extension of helix D from residue 130 to 136, though without direct binding of the heparin pentasaccharide to basic residues in the extended helical region.[254] These changes within the heparin binding site are somehow propagated to the reactive center loop region and are accompanied by expulsion of the hinge residues P15 and P14. However, the side chain of the critical P1 arginine residue points inward to the serpin body in the crystal structures of both native and heparin-bound states, a conformation that would be inappropriate for favorable interaction with proteinase active sites.

Additional solution and X-ray studies have helped to resolve this paradox of the orientation of the P1 side chain and the role of hinge residue expulsion. Thus, in keeping with the expulsion−activation mechanism, it was shown that manipulation of the hinge residue P14 to favor expulsion could promote enhanced reactivity toward factor Xa even in the absence of heparin. Change of P14 from serine to glutamate gave an antithrombin that reacted with factor Xa with greatly increased rate, albeit with increased SI.[255] A time-resolved fluorescence study, monitoring conformational changes at P1 caused by heparin binding showed, however, that the P1 side chain is likely to be solvent, and hence proteinase, accessible in both native and heparin-bound states, so that other interactions outside of this region must be involved in the rate acceleration.[256] This is in accord with mutagenesis studies on reactive center loop variants of antithrombin, which have demonstrated the importance of exosite interactions between factor Xa and heparin-bound antithrombin in enhancing the rate of inhibition.[229] The possible location and nature of this exosite has been suggested by an X-ray structure determination of a P14-fluorescein antithrombin (pdb 1DZG), in which fluorescein covalently bound to an engineered P14 cysteine is too bulky to insert into β-sheet A and instead sits on the surface of antithrombin in the region between the hinge of the reactive center loop and the heparin binding site.[100] This antithrombin species is also activated toward factor Xa, despite having the backbone of the hinge residues P15 and P14 still incorporated into β-sheet A, thereby ruling out loop expulsion per se as the basis for activation. Instead, the presence of the large charged fluorescein moiety on the protein surface alters the electrostatics in a way very similar to that caused by heparin binding. The current working model is therefore that heparin binding induces expulsion of the hinge residues and simultaneously creates an exosite in the region between the hinge of the reactive center loop and the heparin binding site, which has an important electrostatic component. The purpose of expulsion of the hinge region residues may be either to allow contraction of β-sheet A, and hence contribute to the exosite generation, and/or to increase the flexibility of the reactive center loop so that factor Xa can simultaneously bind at its active site to the P1 arginine and at its exosite to the exosite on the serpin body.

While the actual mechanism of linkage between conformational changes in the heparin binding site and expulsion of the hinge region residues is still unresolved, it has been shown to require the ability of helix D to extend. Thus, either introduction of a proline at position 133[257] or removal of one or more residues from the helix extension region[258] results in a partial or complete decoupling of heparin binding from reactive center loop expulsion and activation against factor Xa. A structural model of the nature of the heparin activation mechanism based on these findings is presented in Figure 12. This is an elaboration on an earlier, simpler model that proposed a single step equilibrium between partially loop inserted and loop expelled states that is shifted to the loop-expelled state by heparin binding.[103] The newer model retains these extreme states, but adds an intermediate state in which the reactive center loop has been expelled, but β-sheet A has not yet contracted.

The importance of the specific pentasaccharide sequence has also been extensively explored by solution studies. High affinity heparin binding had previously been shown to result in two-step binding to antithrombin, with the first step involving simple low affinity electrostatic interaction and the second being the induction of a conformational change that greatly enhanced the affinity and thereby locked the antithrombin into the activated state.[227] More recently, rapid kinetic measurements have shown that the three sugars at the nonreducing end (DEF) bind first, and with low affinity. These residues then induce a conformational change in antithrombin that enhances their binding and creates tight binding sites for the reducing end disaccharide (GH). Binding of the reducing end disaccharide then stabilizes the conformation induced by the first three sugars.[259,260]

Whereas the X-ray crystal structure of the antithrombin−heparin pentasaccharide complex gave a very good indication of antithrombin residues that

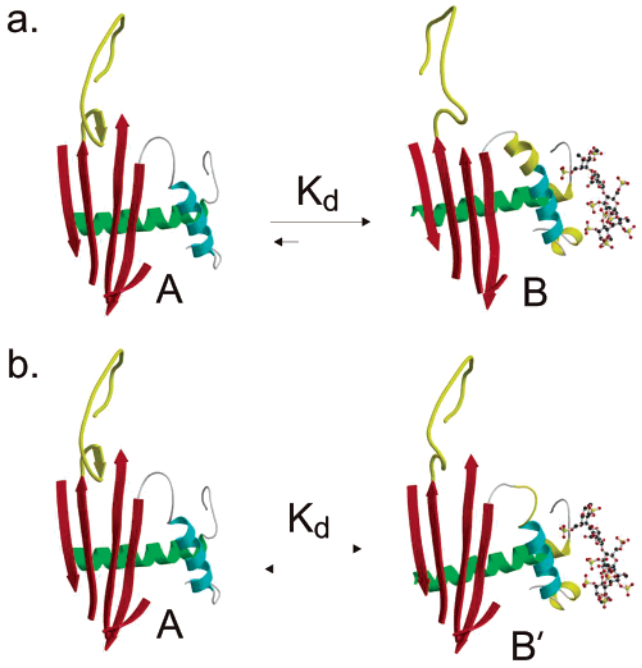

Figure 12. Recent modification of the intermediates proposed to be involved in heparin activation of antithrombin through heparin-induced conformational changes.[257] As in previous models, the native state of antithrombin in the absence of heparin strongly favors partial reactive center loop insertion into β-sheet A, up to and including P14 (state A). Heparin pentasaccharide binding (panel a) causes expulsion of the inserted residues, extension of helix D, and full exposure of the RCL to give the fully active state (state B). It is now proposed that loop expulsion may occur as a separate event preceding extension of helix D and not requiring it. However, helix D extension is required for sheet closure. There would thus be an intermediate state (state B′) in which the top of β-sheet A is still open (panel b). This is a low activity state with intrinsic fluorescence only slightly greater than native antithromin. In cases in which helix D extension is blocked (such as with the deletion variants[254] or the K133P variant[257]), this becomes the end state and represents high affinity heparin binding without significant anti-factor Xa activation. Reproduced from ref 257, copyright (2002) American Society for Biochemistry & Molecular Biology, with permission.

interact with the high affinity heparin pentasaccharide, it did not establish the relative importance of these residues to binding or of which antithrombin residues outside of the pentasaccharide binding site are involved in binding longer chain heparins. A number of recent solution studies have addressed these questions. Mutagenesis studies have established that the single most important residue is lysine 114, which contributes primarily to the conformational change step of the heparin binding process and is also responsible for about 50% of the overall binding energy, albeit in a cooperative manner with other basic residues.[261,262] Lysine 125 is the next most important residue, with replacement by methionine or glutamine causing a reduction in binding energy of 25−33% and also adversely affecting both steps of heparin binding.[263] Arginine 129 has been shown to be similarly important,[264] while arginine 47 is of lesser importance and arginine 46 unimportant.[265] Outside of the pentasaccharide binding site, arginine 132, and lysines 133 and 136, but not lysine 139[254,266] are involved in binding longer chain heparins, despite other work that suggested that lysine 139 might be involved.[267] These conclusions are in accord with sequence conservation and variation within antithrombin sequences from many species, which show absolute conservation of the critical residues but not of arginine 46 or lysine 139.[77]

b. Heparin Cofactor II. The current understanding of the mechanisms of GAG-activation of heparin cofactor II as an inhibitor of thrombin comes largely from solution studies carried out on both wild-type and recombinant variants of the serpin. However, X-ray crystal structures for both native heparin cofactor II and a noncovalent complex with S195A thrombin have recently been determined (codes 1JMJ and 1JMO, respectively), that provide some insight into the structure and role of particular regions of HCII.[104]

HCII differs from antithrombin in a number of important ways. Although both are inhibitors of thrombin, the P1 residue of HCII is leucine rather than the arginine of antithrombin.[268,269] Both dermatan sulfate and heparin are allosteric activators of HCII,[270] whereas only heparin is an effective activator of antithrombin. In addition there is lower affinity of HCII for GAGs compared to antithrombin binding to heparins containing the specific high affinity pentasaccharide. While the region identified as the GAG binding site in HCII shares similar location to that in antithrombin (involving basic residues of the D helix and ones immediately adjacent to it[271−274]), it is believed to be cryptic in the native state of HCII as a result of binding to a negatively charged region within the N-terminal extension region of the molecule.[271,275] This negatively charged region (residues 48−75) is composed of an approximately repeated sequence similar to that found in the C-terminal tail of hirudin, a natural thrombin inhibitor from the leech, and used by hirudin to enhance binding to thrombin by extended interaction with thrombin's basic exosite I.[276,277] The proposed mechanism of conformation change-based activation of HCII, from solution and mutagenesis studies, is that binding of GAG (either heparin or dermatan sulfate) to the helix D basic region of HCII displaces the hirudin-like negatively charged N-terminal tail, which can then bind to exosite I of thrombin and bring the two proteins together (Figure 13). A difference between the effects of dermatan sulfate and heparin is that dermatan sulfate appears to bind only to HCII and contribute to rate acceleration solely by the allosteric mechanism of displacement of the negative N-terminal region, whereas heparin may additionally bridge between HCII and the heparin-binding exosite II of thrombin. However, even in the case of heparin activation with a bridging contribution, the role of allosteric activation is dominant, since mutations of basic residues of the anion exosite II of thrombin, while reducing the affinity of binding of both dermatan sulfate and heparin to thrombin, have no effect on the dermatan sulfate activation of HCII, and only small effects (~7-fold reduction) on the rate enhancement of the heparin activation of HCII.[278]

Evidence in favor of the importance and role of the N-terminal acidic region of HCII in the allosteric

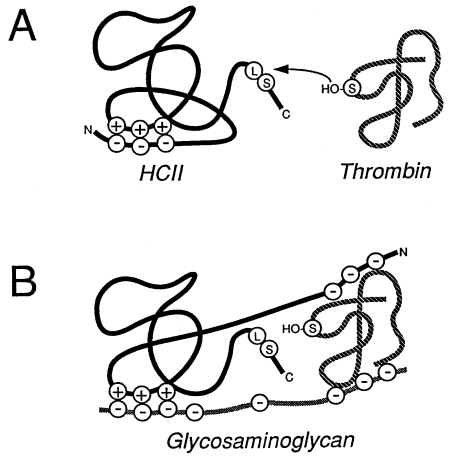

Figure 13. Mechanism of glycosaminoglycan activation of heparin cofactor II proposed by Tollefsen and colleagues.[275] In the native state (panel A), the N-terminal hirudin-like negatively charged tail binds to the positively charged GAG binding site on helix D. The scissile bond in the reactive center loop (Leu-Ser) is only poorly recognized by the thrombin active site. Panel B: GAG binding occurs on helix D and causes the displacement of the negatively charged heparin cofactor II tail, which can then bind to the positively charged exosite I of thrombin, thereby bringing thrombin's active site into appropriate juxtaposition with the scissile Leu-Ser bond of heparin cofactor II. This represents the conformational activation step that is common to binding of both dermatan sulfate and heparin. If the GAG is heparin, an additional bridging interaction can occur for longer chains to exosite II of thrombin. This interaction does not occur for dermatan sulfate. Reproduced from ref 275 with permission. Copyright (1991) American Society for Biochemistry & Molecular Biology.

mechanism is that deletion of this region abolishes most of the GAG rate enhancement, without affecting the ability of HCII to inhibit thrombin in the absence of GAG. In addition, such removal greatly enhances the affinity of HCII for heparin, consistent with its lower accessibility in the native molecule through interaction with the GAG binding site.[271,275] The X-ray structure of the noncovalent complex of S195A thrombin with heparin cofactor II confirms the importance of this region in binding to exosite I of thrombin. What is not yet clear from the X-ray structure of the native molecule is where the N-terminal region is bound, since this region mostly cannot be seen. Whether this is a result of flexibility or altered conformation in the crystal due to monomer−monomer interactions is not known and will require further structural studies.

5.7.3 Regulation of PAI-1 by Vitronectin

Vitronectin is an ~78 kDa multifunctional, multidomain glycoprotein present in the extravascular matrix of many normal tissues and also in high concentration in plasma (~2.5 μM) and the α-granules of platelets. Sequence homology indicates that it is composed of an N-terminal 44-residue somatomedin B domain, followed by two hemopexin repeats, the second of which also contains a heparin binding domain. Through its somatomedin B domain,[279] vitronectin binds tightly to the active form of PAI-1 in the region of the C and E helices and strand 1 of β-sheet A.[280] Because the affinity for latent PAI-1 is ~200-fold lower than for active PAI-1,[281] vitronectin binding serves to increase the half-life for conversion of active PAI-1 to the latent state from ~1−2 h to 4−6 h.[282] Such PAI-1−vitronectin complexes have a number of properties that make them of physiological significance. As an inhibitor both of thrombin[283,284] and activated protein C,[285] the PAI-1/vitronectin complex is much more effective than PAI-1 alone, in each case resulting in an approximately 300-fold rate increase. In the case of the inhibition of activated protein C, the complex is the fastest reported inhibitor, with second-order rate constant of 1.8×10^5 M^{-1} s^{-1}. Whether these rate accelerations are due to vitronectin serving as a template for binding of serpin and proteinase is not known with certainty, but has been suggested.[284] A different effect of complex formation between PAI-1 and vitronectin is on the properties of the latter. The binding site on vitronectin for PAI-1 is immediately adjacent to the RGD site that is used for binding of vitronectin to integrins. Binding of PAI-1 restricts access to the RGD site and so inhibits binding to integrins.[286,287] Since binding of vitronectin to integrins is necessary for smooth muscle cell migration, the presence of PAI-1 can inhibit this by complexing vitronectin. This provides a more subtle rationale for the expression of plasminogen activators at sites of wound healing than initiation of a proteolytic cascade. By reaction of plasminogen activators with vitronectin-complexed PAI-1, vitronectin is freed and made available to bind to integrins. Similar removal of PAI-1 from complex with vitronectin may also result from reaction with thrombin, at a rate accelerated by being in complex with vitronectin. In addition, reactions with thrombin and activated protein C are thought to spare tPA from inhibition and so promote fibrinolysis. This series of interactions thus makes PAI-1 a potentially important regulator of cell migration during wound repair.

5.7.4 Regulation of Protein Z-Dependent Proteinase Inhibitor

Protein Z-dependent proteinase inhibitor[288] (ZPI) is a serpin[289] with tyrosine-serine as the P1−P1′ residues, but one that can nevertheless inhibit both factor Xa and factor XIa in a manner consistent with action as an irreversible suicide substrate inhibitor.[290] Its name derives from the role that protein Z plays in enhancing the rate of reaction of ZPI with factor Xa. Protein Z is a vitamin K-dependent plasma protein with similar structure to the proteinase zymogens, factors VII, IX, X, and protein C, but without any properties as a serine proteinase zymogen.[291] In plasma, protein Z completely complexes all ZPI, without the need for either calcium or phospholipid.[292] However, in the presence of procoagulant phospholipid and calcium, the PZ−ZPI complex inhibits factor Xa at a rate 1000-fold faster than ZPI does alone,[290] suggesting formation of a ternary Xa−PZ−ZPI complex at the phospholipid surface, with PZ acting as a cofactor to present Xa to ZPI. As yet, the nature of the binding interactions or the organization of the ternary complex have not been elucidated. ZPI can also inhibit factor XIa rapidly,

but in a manner that is not dependent on PZ, phospholipid, or calcium.

Elucidation of the possible importance of protein Z in regulating the coagulation response has come from knockout studies in mice, which indicate that protein Z appears to dampen the prothrombotic response. Although unchallenged protein Z null mice have a normal phenotype, those that also have the prothrombotic factor V_{Leiden} gene exhibit a phenotype with much more severe prothrombotic tendency.[293]

5.7.5 Serpin Inactivation by Nontarget Proteinases

The exposed nature of the reactive center loop of serpins is required for serpins to function as proteinase inhibitors. This loop exposure also renders serpins susceptible to attack by nontarget proteinases in a substrate-like manner that results in the inactivation of the serpin, with concomitant conformational changes resulting from loop insertion. Such inactivation may be exploited in vivo, both by endogeneous and exogenous proteinases, for lowering the concentration of functional serpin, and for providing conformationally altered species that may act in a signaling capacity. Metalloproteinases, which serpins cannot inhibit, appear to be particularly important for such proteolytic inactivation.

An early study of the rates of inactivation of three serpins, α_1-PI, antithrombin, and α_1-antichymotrypsin by MMPs 1-, 2-, and 3-, showed that the broad substrate specificity of MMP-3 made it the most effective of the three proteinases in cleaving both α_1-antichymotrypsin and α_1-PI.[294] It has also been shown that MMP-3 can cleave PAI-1 at the P10−P9 and P6−P5 bonds[295] and α_2-antiplasmin at the P3−P2 bond.[296] Although the cleavage positions may differ from one serpin to another, the result in each case is inactivation of the serpin as an inhibitor and generation of a similar loop-inserted state.[297] A possible physiological role for such serpin inactivation is in extracellular matrix remodeling. Indeed, a recent study showed that α_1-PI is a specific substrate for MMP-9 that is cleaved in vivo during dermal−epidermal separation in a mouse model of bullous pemphigod, an autoimmune blistering disease. The dermal−epidermal separation results from increased neutrophil elastase activity as a consequence of MMP-9 inactivation of α_1-PI.[298]

The inactivation of serpins by snake venom metalloproteinases is an example of a different function for such cleavage, the down-regulation of the host's proteinase regulatory system. Thus, it has been shown that serpins of the inflammatory, coagulation, and fibrinolytic pathways can be inactivated by snake venom metalloproteinases.[299,300] As a clever example of evolutionary adaptation to the dangers of such serpin inactivation, the opossum has evolved an α_1-PI with a reactive center loop mutation that renders it resistant to snake venom metalloproteinases, while still allowing it to function as an elastase inhibitor.[301]

Another type of role for nonproductive serpin cleavage is to produce conformationally altered serpins with altered physiological properties. Corticosteroid binding globulin (CBG) is the principal transporter of the glucocorticoids cortisol and corticosterone. Although CBG is a noninhibitory serpin, cleavage within the reactive center loop does result in spontaneous loop insertion, with a resultant 10-fold reduction in the affinity for glucocorticoids for the loop-inserted state, suggesting a role in inflammation resulting from glucocorticoid release at inflammatory loci where neutrophil elastase activity is present.[31,302] Certain cleaved forms of serpins have also been shown to have neutrophil chemotactic properties. These include α_1-PI cleaved at P1−P1′ by elastase,[157] and α_1-antichymotrypsin cleaved by trypsin or *Staphylococcus aureus* metalloproteinase.[158] Finally, reactive center loop-cleaved antithrombin, as well as the structurally related latent form, have been shown to have potent antiangiogenic properties.[156]

5.8 Inhibition of Cysteine Proteinases

There are now a number of well-documented instances of inhibition of cysteine proteinases by serpins. These include inhibition of cathepsins K, L, and S by the serpin squamous cell carcinoma antigen 1,[303,304] inhibition of prohormone thiol proteinase by α_1-antichymotrypsin,[305] and inhibition of members of the caspase family, including caspase 1 (interleukin-1β converting enzyme),[306] caspase 3, and caspase 8 by the viral serpin crmA[307] and caspases 1, 4, and 8 by the human serpin PI9.[308] Both the cathepsins and the caspases differ structurally from serine proteinases that have either the trypsin or subtilisin folds. In addition, the cathepsins and caspases differ markedly from one another; whereas the cathepsins are structurally related to the two-lobe ~24 kDa plant proteinase papain, the caspases are heterodimeric dimers, with each dimer composed of a heavy and light chain, and with a total molecular weight of 60−80 kDa (Figure 14). Whether or not serpins can use the same mechanism that is used to inhibit serine proteinases to inhibit these structurally very different cysteine proteinases is therefore not a trivial question, since it involves not only the effect of replacing oxygen by sulfur in the catalytic apparatus, but also whether the same "crushing" mechanism that is used to compromise the catalytic effectiveness of the serine proteinases could be employed successfully with each of these very different classes of cysteine proteinase.

Given the similarities and differences in the mechanisms of hydrolysis by cysteine and serine proteinases, together with the fundamental features of the serpin suicide substrate inhibition mechanism of reactive center loop insertion, proteinase translocation at the stage of the acyl enzyme intermediate and kinetic trapping of the intermediate through structural distortion of the proteinase active site, there are a number of requirements that would need to be met to allow serpins to inhibit cysteine proteinases by the same mechanism as serine proteinases. These are (i) that the serpin is capable of undergoing the same kind of rapid loop insertion seen with inhibition of serine proteinases, (ii) that it can do this when attached to a cysteine proteinase, (iii) that distortion of the proteinase by compression at the bottom of the serpin is similarly effective in greatly reducing the catalytic operation of the active site, and (iv) that, in the trapped thio-acyl intermediate, rates of deac-

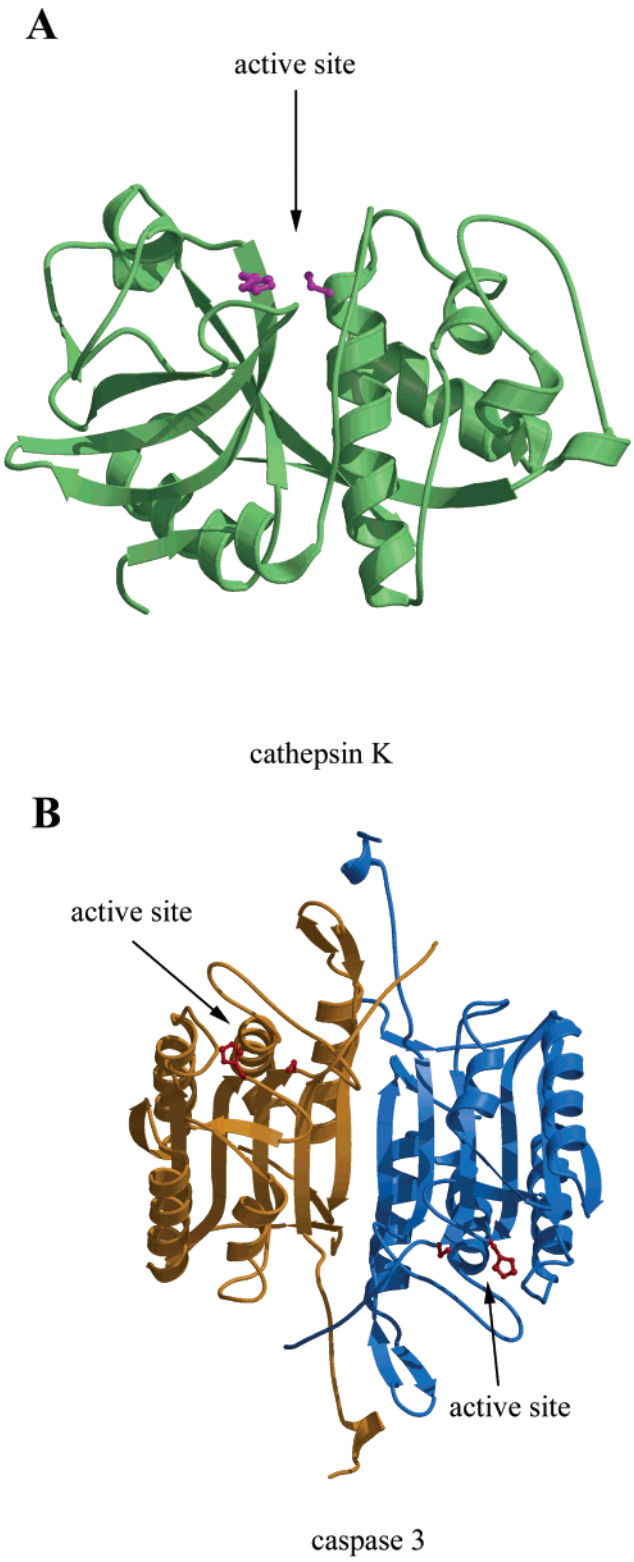

Figure 14. Comparison of the structures of cathepsin K (A) (PDB file 1MEM[501]) and caspase 3 (B) (PDB file 1I3O[502]) showing the accessible surface location of the reactive sites (colored red) of each proteinase, and the surface loops that surround them, and that could provide the fulcrum for pulling the P1 side chain of an attached serpin from the S1 specificity pocket. The caspase is bivalent, and accordingly would require two attached serpin molecules for full inhibition.

ylation are sufficiently slowed to give a kinetically trapped intermediate that would represent an inhibited proteinase on a physiologically relevant time scale.

Of these four criteria, the first is the simplest to address. For crmA, it has been shown that the P14 hinge region residue plays an equivalently critical role in inhibition of the cysteine proteinase caspase 1 as it does in other serpins for inhibition of serine proteinases. Changing this to arginine leads to a large enhancement in the SI for inhibition, consistent with loop insertion being critical to the inhibition process.[309] More direct evidence that the reactive center loop of crmA can insert analogously to those of other serpins came from two recent X-ray structure determinations of reactive center loop-cleaved forms, which have the reactive center loop incorporated into β-sheet A as strand 4A.[93,94] This should not be at all surprising, since three of the four serpins shown to be able to inhibit cysteine proteinases can also inhibit serine proteinases. Thus, crmA is also an inhibitor of granzyme B, α_1-antichymotrypsin is better known as an inhibitor of serine proteinases, and PI9 rapidly inhibits granzyme B in vitro.[425] In addition, SCCA2, which is 92% identical to SCCA1, can be transformed from a predominant serine proteinase inhibitor into a cysteine proteinase inhibitor merely by changing the P1 and P3′ residues.[310]

Concerning the second point of the ability of the proteinase to move over the surface of the serpin during insertion of the reactive center loop, X-ray crystal structures of caspases 1,[311] 3,[312] 7,[313] and 8[314] show similar overall tertiary and quaternary structures for the heterodimeric dimers, with active sites that, while located at the interface of the dissimilar subunits of each heterodimer, are surface accessible (Figure 14). It is therefore unlikely that there would be significantly greater steric hindrance in caspase inhibition than in inhibition of chymotrypsin-like serine proteinases. The same would also be true for papain-like cathepsins such as K, L, and S, which have active sites that lie in the crevice between the left- and right-hand lobes of the molecule, but that are also close to the surface (Figure 14).

Although cathepsins and caspases differ fundamentally from one another in their folds and they in turn differ both from the chymotrypsin-like and subtilisin-like families of serine proteinases, a common feature of each type of structure is an active site located in a cleft or cavity close to the surface and surrounded by loops. In simple mechanical terms, the X-ray structure of the trypsin−α_1-PI complex reveals that the basis for distortion of the proteinase active site is the compression of the loops surrounding the entrance of the active site against the base of the serpin, brought about by the tensioning of the reactive center loop, which is fixed in location at the serpin end by being anchored into β-sheet A and at the proteinase end by covalent attachment through an acyl ester linkage to Ser 195. The loops surrounding the active site thus serve as a fulcrum for pulling serine 195 closer to the base of the serpin and hence distorting the active site. Just as this common physical mechanism can allow for different serine proteinases, and families of serine proteinases, to be inhibited by serpins by a common mechanism, so it

could allow for other proteinase folds to be inhibited analogously. Indeed, in the case of the caspases, the recent elucidation of the mechanism of zymogen activation of caspase 7 by cleavage in a surface loop that is close to the active site, which then adopts a new conformation forming the substrate binding site, emphasizes the importance of loop conformation for catalytic function in caspases.[315]

The last consideration is whether a thiolester intermediate could have sufficient stability to be present in the inhibited complex. This is principally a question of whether the proteinase catalytic activity has been abolished or reduced dramatically, such as occurs in inhibition of serine proteinases, since the thiol ester linkage itself is relatively unreactive toward hydrolysis under physiological conditions.[316] Illustrating this, a recent structure of the 1:1 complex between caspase 8 and the inhibitor p35 showed the presence of a covalent thiolester linkage between the active site cysteine of the caspase and a residue of the inhibitor at the site of proteolytic cleavage.[317] Similarly, thiolesters are present as mechanistically essential components of complement proteins C3 and C4[318] and of the broad-spectrum proteinase inhibitor α_2-macroglobulin.[319] In α_2-macroglobulin, the thiol ester has long-term stability against hydrolysis, even in the unactivated state, but has limited accessibility and reactivity toward small nitrogen nucleophiles.[320] Similarly, the thiol esters in C3 and C4 are stable against hydrolysis unless the protein is activated by proteolysis, after which the thiol ester is itself activated by adjacent groups.[321,322] The only experimental evidence to date for formation of a thioacyl intermediate in a serpin/cysteine proteinase reaction is that of antithrombin with papain.[216] In this study, an excess of antithrombin was reacted with papain. Under these conditions, a high molecular weight SDS-stable complex was formed at the same time as the free papain concentration decreased. However, careful analysis showed that the reaction, while involving formation of a thioacyl intermediate, as expected for a serpin-type reaction with a cysteine proteinase, was complicated by the consequences of greatly enhanced autoproteolytic susceptibility of the antithrombin-complexed papain. The result was that the papain trapped in the covalent thioacyl papain−antithrombin complex was rapidly cleaved to small peptides. Presumably as a result of this extensive cleavage, the thioester linkage became solvent exposed and could react with nitrogen nucleophiles present on other antithrombin molecules. This gave rise to covalent antithrombin dimers (the high molecular weight species visible by SDS−PAGE), in which the C-terminus of the heavy chain of the cleaved antithrombin (the molecule that was earlier involved in the thioacyl linkage to papain) was linked via an amide or ester bond to an intact antithrombin molecule. This type of protein−protein cross-linking, resulting from attack of a protein nucleophile on a thioacyl carbonyl, is the basis for proteinase cross-linking to activated α_2-macroglobulin[323] and complement components C3 and C4.[322] Although this study on the antithrombin/papain reaction was thus rather more complex than expected, it did illustrate both that a thioacyl intermediate was formed in the expected way and that the initially complexed papain must have been subjected to major conformational distortion (as expected from the known serpin mechanism with serine proteinases) to explain the rapid autoproteolytic degradation that papain is otherwise not susceptible to. That no stable papain−antithrombin complex was formed might be due to the position of attack by papain within the antithrombin reactive center loop, which occurred at either P2−P1 or P2′−P3′. In neither case would this result in fulfillment of the usual length requirement of 17 residues for the portion of the reactive center loop inserted into β-sheet A, but instead would be either one residue shorter or two residues longer. As discussed in section 5.5.3, such changes in length are usually associated with failure to trap the acyl enzyme intermediate of serine proteinase reactions in a stable complex with serpins.

An additional point that needs to be discussed in the context of possible differences between the inhibition of cysteine proteinases and serine proteinases by serpins is the behavior of the inhibited complexes on SDS−polyacrylamide gels. The demonstration of an SDS-stable high molecular weight complex has become something of a hallmark, and hence diagnostic, for operation of the suicide substrate loop-insertion mechanism of serpin inhibition for serine proteinases. Indeed, the early demonstration that this high molecular weight band corresponded to a complex between the serpin and proteinase via an acyl linkage is consonant with the finding of an acyl linkage between the P1 carbonyl and the proteinase serine 195 γO in the crystal structure of the non-denatured trypsin−α_1-proteinase inhibitor complex. By implication, the failure to detect an equivalent SDS-stable high molecular weight species for complexes of serpins with cysteine proteinases has been taken as an indicator of a fundamental difference in mechanism or in stability *of the intact complex*. However, it is an oversimplification, even in the case of serine proteinase inhibition by serpins, to equate the presence of an SDS-stable high molecular weight complex with attainment of the *same* final structure as was found in the single X-ray structure determination. All this really demonstrates is that an acyl intermediate was present at the time of denaturation, and that the deacylation rate must have been compromised sufficiently in the complex, at that point, that there was insufficient time for deacylation to occur before complete SDS-denaturation had occurred. This is because, once denatured, the acyl intermediate itself is slow to hydrolyze under the conditions of running the gel. In this regard, it should be noted that a study that claimed to follow the kinetics of *final* covalent complex formation between chymotrypsin and α_1-antichymotrypsin by careful, rapid quench SDS−PAGE analysis of the kinetics of appearance of high molecular weight covalent acyl intermediate complexes, was in fact almost certainly only following formation of the acyl intermediate.[324]

In contrast to acyl intermediates, which react slowly at physiological pH with both oxygen and nitrogen nucleophiles, a thioacyl intermediate is quite

reactive in the presence of nitrogen nucleophiles. It is therefore much more difficult to effect a fast enough denaturation and gel electrophoretic analysis of a thioacyl intermediate, especially from a complex in which there may be some residual catalytic effectiveness within the proteinase, giving a much shorter time window for denaturation and analysis before deacylation has occurred. In addition, the usual buffer conditions for running the SDS−PAGE would provide nitrogen nucleophiles that could deacylate the denatured thioacyl intermediate. This point was addressed in the case of the caspase 8-p35 complex, where the X-ray structure clearly showed the presence of a covalent thio-acyl intermediate. SDS−PAGE under normal conditions showed no stable covalent complex between proteinase and inhibitor. However, by lowering the pH and using a lower temperature for denaturation, it was possible to demonstrate the presence of just such a linkage by SDS−PAGE, albeit at less than the expected stoichiometry.[317]

In conclusion, while there is still a shortage of firm structural evidence for the operation of the same mechanism of inhibition of cysteine proteinases by serpins, of the kind that is now available for serine proteinase inhibition, it seems that the same mechanism could accommodate all of the conditions necessary for effective inhibition of this second class of proteinase, including the formation and stabilization, in the complex, of a thioacyl intermediate.

5.9 Inhibition by Reversible Complex Formation

Whereas all consideration of inhibition and complex formation by serpins, whether of serine or cysteine proteinase, has so far dealt with covalent, effectively irreversible complex formation, there are several reports that suggest that serpins can not only employ that unusual mechanism, but can also inhibit by readily reversible noncovalent complex formation in a manner much more like that of Kunitz and similar families of inhibitor. However, from the small number of such reports compared to the very large number of examples of covalent inhibition, it seems likely that the noncovalent mechanism can only operate under exceptional circumstances.

The three most convincing examples of such inhibition are of reversibility of chymotrypsin complex formation with α_2-antiplasmin,[325] of single chain urokinase (scuPA) inhibition by protein C inhibitor,[326] and of factor Xa inhibition by an unusual mosquito serpin, AFXa.[69] In the first study, complexes of chymotrypsin or trypsin were generated with the serpin α_2-antiplasmin by incubation for 1 min, followed by incubation with α_2-macroglobulin to trap any active proteinase that might be released from the serpin/proteinase complex. A dissociation rate constant of 5.5×10^{-5} s^{-1} was found for the chymotrypsin/α_2-antiplasmin complex, and the free α_2-antiplasmin that was formed was capable of further inhibiting added proteinase, indicating that it was still functional rather than cleaved. While this is an interesting study, it involves a nonphysiological pair. In the case of scuPA complex formation with PAI-3 (protein C inhibitor), this is a much more physiologically important interaction. ScuPA, although the zymogen form of two-chain urokinase (tcuPA), is one of a small number of zymogens that possesses measurable, low catalytic activity. However, two puzzles have been (i) scuPA in biological fluids has no measurable activity against plasminogen and (ii) plasmin, formed by activation of plasminogen, is itself required to activate scuPA into tcuPA. It was found that 1:1 scuPA/PAI-3 complexes were reversibly formed in the presence of U937 cells expressing surface urokinase receptor, uPAR. Since it has been shown that uPAR-bound scuPA is 2−3 orders of magnitude more reactive toward plasminogen than is free scuPA, this suggests a role for reversible scuPA/PAI-3 complex formation as an inactive storage form of scuPA. Furthermore, in the uPAR-bound state it was found that scuPA no longer bound to PAI-3 and could therefore no longer be inhibited by it. The proposal is therefore that when plasminogen activation is needed in vivo, cells express uPAR, which scavenges scuPA from reversible noncovalent complexes with PAI-3 to give an active proteinase for activation of plasminogen. In contrast, tcuPA bound to uPAR can be inhibited by PAI-3 or PAI-1. The study therefore suggested that formation of tcuPA represented a self-limiting capacity for plasminogen activation, in that conversion of uPAR-bound scuPA into uPAR-bound tcuPA allowed the proteinase to be both inhibited by PAI and to then be cleared by binding to LRP and being internalized. Although the nature of binding in the noncovalent complex of PAI-3 and scuPA has yet to be firmly established, it is known to be highly dependent on the P1 residue, and also to result in perturbation of a P2′-attached exogenous fluorophore, both of which suggest involvement of the reactive center loop residues.

The third study is of reversible, tight, noncovalent, but also noncompetitive, binding of mammalian factor Xa to the *anticoagulant factor Xa* (*AFXa*) gene product of the yellow fever mosquito *Aedes aegypti*.[69] Complex formation may be a 1:1 interaction, but does not result in formation of an SDS-stable high molecular weight complex. AFXa is a serpin, but one with a reactive center loop that is shorter than normal (three residues shorter than usual on the side N-terminal to the potential reactive bond) and with an amino acid sequence in the hinge region that would be inimical to facile loop insertion. The kinetic analysis of inhibition furthermore suggested binding of AFXa to an exosite on factor Xa, rather than the substrate specificity subsites. Nevertheless, the reactive center loop does contain a sequence, VDRR, that is somewhat similar to the IEGR preference of factor Xa, so that binding may involve contacts both with an exosite on AFXa and with part of its reactive center loop, not in a manner conducive to substrate-like cleavage, but perhaps instead one more like a canonical-type proteinase inhibitor.

In at least two, and possibly all three, of the above three examples there is probably no need to invoke a new pathway for interaction between serpin and proteinase from that depicted in Figure 6 and that is the basis for covalent inhibition. Since ultimate covalent inhibition requires initial formation of a

noncovalent Michaelis-like complex, it is conceivable that examples of reversible noncovalent inhibition represent instances in which there are important additional stabilizing interactions in the initial complex that would be lost either upon cleavage of the peptide bond or upon initiation of loop insertion, or else an intrinsically slow reaction rate for formation of the acyl intermediate, such as with the low activity zymogen scuPA. Evidence for residues within the reactive center loop stabilizing the noncovalent complex comes from a study on a P1′ Ser → Leu mutant of antithrombin, in which the P1′ leucine slowed the conversion of the initial noncovalent complex into the covalent intermediate by 190-fold.[327] If, in addition, exosite interactions involving the body of the serpin and an exosite on the proteinase were also present and of sufficient strength, the stabilization of the nonloop inserted form might be enhanced sufficiently to make that species either kinetically or thermodynamically stable, since loop insertion would probably involve disruption of such exosite interactions and loss of that binding energy. That there is energy enough to stabilize such a complex is evident from studies on some Michaelis-like serpin complexes with anhydroproteinases. Thus, whereas anhydrotrypsin binds to PAI-1 with K_D of 0.23 μM, S195A t-PA binds with K_D of 1.3 nM, resulting from additional exosite interactions.[162] Although this does not result in noncovalent complex formation between active tPA and PAI-1, it does cause reversible acylation by impeding release of the cleaved N-terminus of the reactive center loop from the active site of tPA.[328]

6. Clearance and Signaling

6.1 Receptor-Mediated Clearance

In humans, the half-life for serpin−proteinase complexes in the circulation is very much shorter than for native serpins, with the former being only a few hours compared with over 24 h for the latter.[329,330] Studies on the clearance of radiolabeled human serpins injected into mice showed that, although clearance was generally much faster than for the same proteins in humans, there was still ~10-fold faster clearance of serpin−proteinase complexes than of native serpins.[331] In addition, these studies, which examined clearance of several serpin−proteinase pairs, including thrombin−antithrombin, heparin cofactor II−thrombin, α_1-PI−trypsin and α_1-antichymotrypsin−cathepsin G, showed that a common receptor appeared to be involved, at least for these serpins.[332,333] In the case of α_2-antiplasmin−plasmin complexes, binding may involve a separate receptor, or the same receptor, but with much lower affinity.[334]

The clearance of many serpin−proteinase complexes by the same receptor, and the specific recognition of complexes over native or even cleaved serpin, led to the idea of a common epitope on serpins that was cryptic in the native state and that became exposed only in complexes. Perlmutter and colleagues identified a region in the C-terminal region of α_1-PI that appeared to compete with various serpin−proteinase complexes for binding to human HepG2 cells[335] and on this basis proposed that a nearly conserved hydrophobic pentapeptide in this C-terminal region of many serpins was the common binding site.[336] The pentapeptide, from P12′ to P16′, has the sequence FVFLM in α_1-PI and shows very high conservation among inhibitory serpins, being FLMII in α_1-antichymotrypsin, FLFLI in heparin cofactor II, and FLVFI in antithrombin. This proposal was criticized on the basis of the known structure of cleaved α_1-PI, which showed that this hydrophobic region was largely buried in the cleaved serpin.[337] It was later shown that single residue mutation of each of these positions to alanine in heparin cofactor II had no effect on binding to receptor and on clearance of HCII−proteinase complexes.[338] Knowing now that the serpin conformation in the covalent complex is almost identical to that in the cleaved state, so that the peptide is also buried in the complex, has made this specific proposal for a serpin-receptor binding site quite untenable.

6.2 LRP as the Principal Clearance Receptor

The first demonstration that LRP is involved in clearance of proteinase−proteinase inhibitor complexes came from attempts to identify the receptor responsible for binding and clearing complexes of proteinase with α_2-macroglobulin, rather than from studies on serpin−proteinase complexes. These studies demonstrated that the α_2-macroglobulin receptor, which specifically recognizes the conformationally altered form of α_2-macroglobulin, was the same protein as the low-density lipoprotein receptor-related protein (LRP).[339,340] It was soon shown that LRP was also capable of binding and clearing from circulation complexes of uPA and tPA with the serpin PAI-1.[341,342] It has subsequently been shown that LRP binds to the α_1-PI−elastase complex,[343] complexes of antithrombin and heparin cofactor II with thrombin,[344] uPA complexes with protease nexin 1, C1-inhibitor, and protein C inhibitor, and thrombin complexes with protease nexin 1, and protein C inhibitor,[345] C1-inhibitor−C1s complexes,[346] and factor XIa−protease nexin 1 complexes.[347] Thus, evidence for the importance of LRP in clearance of serpin−proteinase complexes is now overwhelming. These interactions are summarized in Table 11.

6.3 Structure of LRP and Other LDL Receptor Family Members

LRP is a member of the LDL receptor family of proteins, all of which are transmembrane proteins composed of several repeated domains (Figure 15), and are responsible for binding a range of ligands, many associated with lipoprotein metabolism and proteinase metabolism. LRP is one of the largest members of the family, being about three times the size of the LDL receptor. It is synthesized as a 600 kDa precursor that is cleaved in the trans-Golgi by furin to a 515 kDa α-chain and an 85 kDa β-chain,[348] which associate noncovalently, with the α-chain on the extracellular side of the membrane. The β-chain contains the single transmembrane domain. The α-chain contains 31 copies of the domain termed

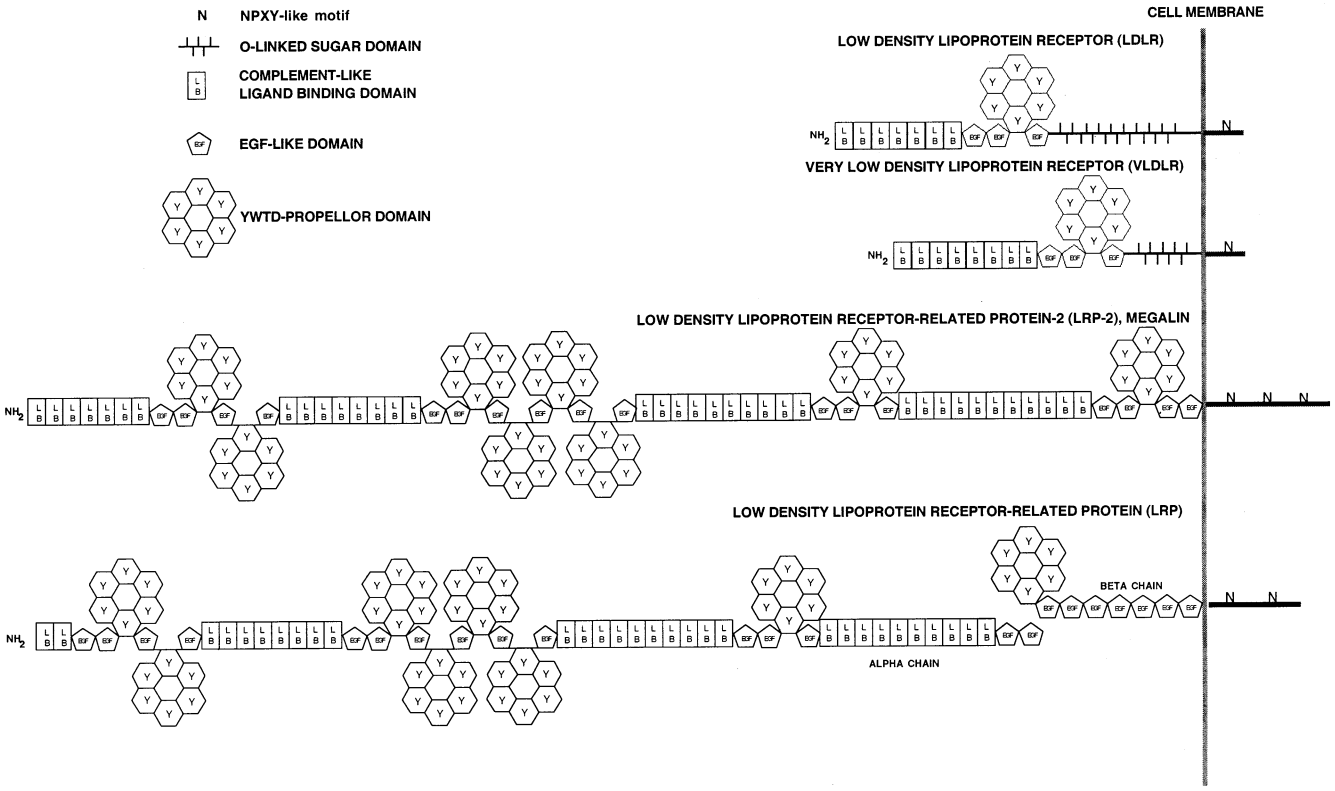

Figure 15. Schematic depiction of members of the LDL receptor family of proteins, showing the location and multiplicity of the repeated structural repeats present in these receptors.

Table 11. Relative Binding Affinities[a] of Different LDLR Family Receptors for Different Serpin−Proteinase Complexes

complex	LRP	LRP2 (megalin)	VLDLR	ref
α_1-PI−trypsin	0.07[b]			344
α_1-PI−elastase	binds & clears	binds & clears		343
α_1-ACT−cathepsin G	no uptake	binds & clears		343
PCI−thrombin	0.2		0.6	345
PCI−uPA	1.9		0.3	345
PN1−thrombin	0.2[c]		9	345
PN1−uPA	2.7		15	345
antithrombin−thrombin	0.14		<0.01	345
antithrombin−uPA	0.02		<0.01	345
heparin cofactor II−thrombin	<0.02		<0.01	345
PAI-1−uPA	1.9		1.0	345
PAI-1−tPA	2.0		0.7	345
PAI-1−thrombin	0.04[d]	0.01[d]		363
C1 inhibitor−C1s	binds & clears			346
C1 inhibitor−uPA	0.08		0.01	345

[a] Almost all reported measurements are of relative binding rather than direct determination of K_d. The values reported here are therefore affinities relative to that of PAI-1:uPA, taken as one. The higher the number, the tighter the binding. K_d values have been reported for protease nexin 1 complexes with uPA binding to VLDLR and LRP as 0.14 and 0.8 nM, respectively. [b] Interpolated to data from ref 345 using relative affinity for LRP of complex of thrombin-antithrombin. [c] Clearance reported to be strongly dependent on heparin.[391] [d] Interpolated to data from ref 334 by use of reported K_Ds for PAI-1-thrombin to LRP and LRP2.

"ligand binding" (LB) or "complement-like repeat" (CR). These are ~40 residue domains with six conserved cysteines that form three disulfides and a calcium binding site with almost completely conserved carboxyl ligands. The CR domains of LRP are organized into four clusters of 2, 8, 10, and 11 repeats. Such clusters are thought to be the principal binding regions for the various protein ligands within the whole family of receptors.[349,350] In LRP, these ligand binding regions are flanked by some of the 22 EGF-like repeats, and further separated by intervening groups of six "YWTD" domains, that have recently been shown to form a tightly associated 6-propellor structure.[351,352] The X-ray structure of the YWTD−EGF domain pair demonstrated a close interaction between these two motifs. On the intracellular side, the ~110 residue C-terminal tail contains two NPXY "internalization" motifs.[353] The largest member of the family so far identified is megalin, also known as gp330 or LRP2. The VLDL, LDL, and apoE receptors are much smaller, single chain species that contain only a single cluster of CR domains and a single copy of the 6-propellor YWTD cluster.

Both X-ray[354,355] and NMR[356−360] structures have been reported for individual and tandem CR domains, from both LRP[355,358,360] and LDLR.[354,356,357,359] All show the same pattern of disulfide bonds, a similar two-lobe fold, the same type of calcium coordination in the C-terminal lobe, but different surface electrostatics in each case, resulting from lack of conservation of most residues, other than those involved in calcium coordination or disulfide formation. Studies on a tandem domain have shown that each domain has complete freedom of movement relative to the other domain.[359,361]

6.4 Specificity of Binding

6.4.1 LRP versus LRP2 and VLDLR

Although LRP has been shown to have the largest repertoire of protein ligands of all the LDL receptor family members, including apolipoproteins and lipases, α_2-macroglobulin−proteinase complexes, many serpin−proteinase complexes, some proteinases, certain matrix proteins, and miscellaneous other ligands including lactoferrin and rhinovirus,[362] there is considerable overlap in specificity of binding of serpin−proteinase complexes with LRP2 and the VLDL receptor. A comparative study of complex binding to LRP and VLDLR examined the relative affinities of different serpin−proteinase complexes for each of the two receptors. Using microtiter plates coated with purified receptor and purified proteinases, differences were found not only in the range of complexes bound by each receptor, but also in the rank order of affinities of different complexes for each receptor.[345] VLDLR showed over a 1000-fold difference between the tightest complex bound (uPA−protease nexin 1) and the weakest (uPA−antithrombin and thrombin-HCII), whereas LRP showed over a 100-fold difference between the tightest (uPA−protease nexin 1) and the weakest (uPA−antithrombin and thrombin−HCII). In addition, it was found that for the tightest binding complex, uPA−protease nexin 1, the affinity to the VLDLR was higher ($K_d \sim 0.14$ nM) than to LRP ($K_d \sim 0.8$ nM). Another study that looked at internalization and clearance rates by mouse fibroblasts found very large differences and suggested that this also reflected differences in affinity, with uPA−PAI-1 being cleared much more effectively than antithrombin−thrombin, HCII−thrombin or α_1-PI−trypsin complexes.[344] Comparative rates of clearance of radiolabeled thrombin in complex with PAI-1, antithrombin, HCII, and α_1-PI by LRP and LRP2 have been examined using anti-LRP and anti-LRP2 antibodies and the universal competitive ligand RAP (receptor associated protein) to demonstrate clearance by these receptors.[363] PAI-1, further enhanced by being complexed to vitronectin, was the most effective serpin for this clearance, via both LRP and LRP2, though binding to LRP2 was about 4-fold weaker. Finally, a comparison of cellular uptake and degradation of α_1-PI−HNE and α_1-antichymotrypsin−cathepsin G complexes by cells that expressed either LRP or LRP2 showed an inability of LRP to take up and degrade the α_1-antichymotrypsin−cathepsin G complex, whereas LRP2 was effective against both.[343]

Because of the relative paucity of detailed comparative studies of all three receptors with a wide enough range of serpin−proteinase combinations, it is hard to make definitive generalizations about the detailed specificities of the receptors. It is clear, however, that each receptor has a different range of serpin−proteinase specificities, with LRP probably having the largest range. Of the different types of complex examined, it is usually those with uPA as proteinase or protease nexin 1 or PAI-1 as serpin that seem to have the highest affinities for receptor, though there are differences for different receptors. Thus, while both VLDLR and LRP have the highest affinity for uPA−protease nexin 1 complexes, LRP2 binds α_1-antichymotrypsin−cathepsin G tightest (Table 11). In this context, it should be noted that the in vivo distribution of these receptors is cell and tissue specific, so that affinity and specificity may be tailored to the need in a given cell and environment. LRP is known to be abundant in hepatocytes and many other cell types,[364] LRP2 in kidney cells,[365] and VLDLR in skeletal muscle, heart, adipose tissue and brain, but to be present at low levels in the liver.[366]

6.4.2 Role of Serpin and Proteinase Moieties in Binding

Some of the earliest studies on clearance of serpins from circulation showed that native and cleaved forms had much longer half-lives than covalent complexes.[367] This implies tighter binding of the complex, which might arise either from the generation of a unique epitope within the serpin or else of a binding site with elements composed of both serpin and proteinase. A study on PAI-1 complexes with high and low molecular weight uPA, tPA and trypsin concluded that, while high affinity binding required complex formation, it was independent of the nature of the proteinase, since different PAI-1 complexes could cross-compete with one another, implying that the high affinity receptor binding epitope resided in the serpin alone.[368] Given the structure of the covalent α_1-PI−trypsin complex compared to that of cleaved α_1-PI, this is hard to understand, since the differences in the serpin moiety consist of sites that are obscured by the proteinase rather than revealed. It should, however, be realized that this study showed that all other uncomplexed forms of PAI-1 *could* compete for binding of covalent complex, albeit with an affinity 2−3 orders of magnitude lower (K_d in the 0.1−1 μM range). In addition, it is clear from other studies that the proteinase can make a significant difference in binding affinity. Thus, the uPA−protein C inhibitor complex binds 100-fold more tightly than the uPA−antithrombin complex to LRP, whereas the thrombin−protein C inhibitor complex binds with similar affinity as the thrombin−antithrombin complex.[345] Studies on in vivo clearance of C1-inhibitor complexes found a 2-fold difference in half-lives as a function of the proteinase in complex with the serpin.[369,370]

A possible explanation for these apparently contradictory studies is that both serpin and proteinase moieties can have binding sites on the receptor that contribute to the overall affinity. Each may contribute to the total in an additive manner, with the serpin probably contributing in the micromolar range and the proteinase contibuting substantially less. Nevertheless, together they could have a very tight interaction with the receptor that depends on the serpin being in complex with proteinase, in a way that will depend on the nature of the proteinase. However, since the proposed proteinase contribution to the binding interaction could be relatively weak (perhaps in the 0.1−10 mM range), it can afford to be relatively nonspecific and therefore nondemanding as far as finding a binding site on LRP in close

enough proximity to the higher affinity site for the serpin, and would show no more than a 100-fold variation as a function of proteinase for most nonspecific proteinase binding. Of course, where there is a higher affinity binding site for the proteinase itself, such as has been shown for uPA,[371] overall affinity for the complex with serpin would be even higher. This is significant, since, as noted above, serpin complexes with the plasminogen activators usually bind with much higher affinity than complexes with other proteinases.

6.4.3 Residues Involved in Binding

Relatively little has been firmly established at the molecular level with respect either to the structural requirements on LRP for binding protein ligands or on the ligands themselves for binding to the receptor. Although there is a requirement for calcium for protein ligand binding to LRP,[372,373] this can now be understood in terms of the internal calcium binding sites of each CR domain, which are needed to maintain a rigid conformation for the domain,[374] rather than the calcium acting as a metal ion bridge between the two proteins.

It has been proposed, however, that binding of various protein ligands to LDL receptor family members may involve basic residues on the ligand interacting with acidic regions on the receptor. Mutagenesis studies on apoE, in which basic residues were changed to alanine, resulted in compromised binding.[375] The X-ray structure of the receptor-binding region of apoE revealed the location of these critical basic residues to be on the outer face of a helical region.[376] Similarly, it has been shown that basic residues within the receptor-binding region of α_2-macroglobulin are necessary for binding.[377] On the basis of the possibility of such a general mechanism of receptor binding, selected mutations of basic residues in PAI-1 were made and shown to result in 3−20-fold weaker binding of uPA−PAI-1 complexes to purified LRP and VLDLR and to reduce levels of endocytosis by both receptors by 40−50%.[368,378] The mutated residues cluster to one face of PAI-1, composed of parts of strand 1 of β-sheet A and helix D. One study that may argue against a general basic−acidic mechanism of binding is on complexes of thrombin−protease nexin 1. Peptide screening identified a peptide that could greatly reduce complex internalization.[379] Mutagenesis of protease nexin 1, based on these peptide results, identified an adjacent histidine−aspartate pair as critical for internalization, though not for cell surface binding.[380] It therefore remains to be seen whether there is a common involvement of basic residues in binding of other serpins to LRP.

Concerning requirements on the receptor, it has been proposed, based on studies in which individual CR regions of LDLR were deleted, that a ligand binding site may be composed of interactions with more than one CR region and that by using different combinations of domains a wide range of binding sites can be generated.[381] More recently, similar behavior has been observed for VLDLR. Absence of the third CR domain results in lowered affinity for RAP, but not for two serpin−proteinase complexes, despite the fact that both complexes competed for RAP binding.[382] There have also been studies that have localized binding of particular protein ligands to different CR domains of LRP. A study that examined each tandem pair of CR domains within cluster II of LRP (as fusion proteins) localized binding of uPA-PAI-1 to the pair of domains CR5−CR6,[383] though with affinity of only $\sim$1 μM, compared with 0.4 nM for binding to the intact receptor,[384] suggesting that a significant part of the full binding site was still missing. Limited mutagenesis of the CR5−CR6 pair was also carried out on the two residues in each domain that coordinate the calcium through their backbone carbonyl rather than their side chains (a tryptophan and an aspartate in each). Loss of binding of the tandem construct to uPA-PAI-1 resulted. Given the structure of related domains CR3[360] and CR8,[358] the aspartate-to-asparagine change is unlikely to perturb the structure and so may implicate a negative site in each domain as required for binding. The tryptophan-to-serine mutation may, however, be less well tolerated in these small domains that contain only minimal hydrophobic cores. Full resolution of the nature of ligand binding may need to await a structure determination of a ternary complex between serpin−proteinase binary complex and a cluster of CR domains.

6.5 Internalization and Degradation Mechanism

As with metabolism of other ligands that are internalized and degraded as a result of binding to LDL receptor family receptors, there are three distinct phases to the metabolism of serpin−proteinase complexes. The first is binding to the receptor, the second is internalization and the third is release of the serpin−proteinase complex for internal degradation, with concomitant return of the receptor to the surface. This mechanism was worked out in detail for LDL receptor-mediated internalization of cholesterol,[385] and involves spontaneous and continuous clustering of the receptor into clathrin-coated pits independent of bound ligand, followed by internalization, release of the ligand from the receptor as a result of a pH drop, degradation of the ligand, and recycling of the receptor to the cell surface. The demonstration that different serpin−proteinase ligands may bind to one or more of the receptors LRP, LRP2, and VLDLR raises the possibility that the binding, internalization, and release mechanisms may be distinct for each receptor.

Whereas many studies described above, and aimed at examining the specificity of different receptors for particular serpin−proteinase complexes, have used purified receptor and ligand components, it is clear that co-receptors can play important roles in binding of some serpin−proteinase complexes. In the case of uPA−PAI-1 and uPA−protease nexin 1 complexes, binding to the uPA receptor uPAR is required for LRP-mediated[386,387] or VLDLR-mediated[388] internalization of these serpin−proteinase complexes. This probably involves formation of a tetramolecular complex between uPAR and LRP/VLDLR, bridged by the serpin-uPA complex, which is then internalized.

Once the serpin−proteinase complex has been removed intracellularly, uPAR recycles to the cell surface.[389,390] LRP thus serves not simply to clear uPA−serpin complexes, but to regulate cell surface uPAR activity. For internalization of thrombin−protease nexin 1 complexes, it has been shown that cell surface-associated heparin chains are required for efficient binding and internalization,[391] raising the possibility that, for serpin−proteinase complexes in which either or both component can bind heparin, such surface binding to heparin-containing species may serve to concentrate the serpin−proteinase complex and ensure higher saturation of the LRP binding site.

Although it has been shown for the LDL receptor that internalization of the receptor via clathrin-coated pits, with or without bound ligand, involves the cytoplasmic NPXY motif,[353] the presence of 1, 2, and 3 copies of this motif in VLDLR, LRP, and LRP2, respectively, does not ensure that the same internalization mechanism operates in each case. Thus, it has recently been shown that it is the overlapping YXXL motif in the cytoplasmic tail that is more important for internalization in the case of LRP.[392] In addition, there are reports that both LRP-mediated[393] and VLDLR-mediated[394] ligand internalization depend on intracellular serine phosphorylation by a cAMP-dependent protein kinase. In the case of internalization via LRP, it has also been shown that the LRP cytoplasmic domain interacts directly with a GTP-binding protein,[395] thus making LRP a signaling receptor. In light of this, it is likely that the so-called α_2-macroglobulin "signaling receptor", operationally defined and championed by Pizzo and colleagues[396] as a receptor distinct from LRP, is in fact one and the same as LRP.

Some light has recently been shed on how ligand binding to monomeric LRP might stimulate intracellular signaling by demonstrating the proximity of LRP to another transmembrane protein, amyloid β precursor protein, through intracellular binding to the adaptor protein Fe65.[397] A more extensive yeast two-hybrid analysis has been carried out on potential binding of a range of adaptor proteins to the cytoplasmic tails of all LDL receptor family members, with the finding that LRP and LRP2 can interact with Dab-1, PSD-95, CAPON, JIP-1, and -2, and SEMCAP-1, which, in turn, can link these receptors to many cellular signaling pathways.[398]

Once the receptor−serpin−proteinase complex has been internalized and the serpin−proteinase complex dissociated from the receptor, the fates of the two species are different, with the former directed to degradation and the latter recycled to the cell surface for reuse. An important question, however, is what brings about dissociation of the serpin−proteinase complex from the receptor. While the drop in pH to ~5 caused by fusion of the endocytosed vesicle containing the internalized receptor−ligand complexes with endosomes is likely to be crucial, the mechanism of this pH-dependent dissociation has yet to be worked out. The discovery that the CR domains contain an internal calcium binding site,[354] and that calcium is required to lock the conformation of the domains,[374] led to the suggestion that the pH drop might cause calcium dissociation from these repeats and hence ligand dissociation.[354,399] However, it is clear from calcium binding studies at pHs 7.4 and 5 that calcium affinity does not drop by enough in this pH range to explain such ligand dissociation.[355] Indeed, the pH-dependence of ligand binding has been shown to be even more pronounced. However, it has been shown both for LRP[400] and VLDLR[401] that the flanking EGF domains are critical for ligand release, though not for binding. The pH-dependence of binding may thus involve a pH-dependent conformational change involving the association of the CR clusters, where the ligand binds, with the YWTD propellor domains, mediated by the intervening EGF domain.[351] Finally, there is evidence, at least in the case of uPA−protease nexin-1 complexes that, once internalized, binding to endosomal heparins may be needed for retention and degradation.[402]

6.6 Signaling Mechanisms

In the preceding section, some evidence was presented for how LRP, LRP2, and VLDLR may act as signaling receptors following ligand binding, through binding of their cytoplasmic tails to intracellular adaptor proteins. As noted, specific details of linkage to such activation pathways from the initial serpin−proteinase ligand binding step are few. However, the linkage through the adaptor proteins to signaling pathways involved in cytoskeletal rearrangement, proliferation, cell adhesion, and apoptosis, among others, provides a plethora of potential ways in which serpin−proteinase complexes may serve as signaling molecules.[398] There are in the literature a number of examples of downstream responses from generation of serpin−proteinase complexes that might involve such pathways. α_1-Antichymotrypsin−cathepsin G complexes have been shown to stimulate production of interleukin 6 by human lung fibroblasts, while conditioned media from these cells could stimulate production of α_1-antichymotrypsin and other acute phase response proteins.[403] α_1-Antichymotrypsin−chymotrypsin complexes were also shown to inhibit the activation of NADPH oxidase that normally results from stimulation of neutrophils with the potent chemotactic agent fMet-Leu-Phe.[404] α_1-PI-neutrophil elastase complexes have been shown to have neutrophil chemotactic properties associated with a region in the C-terminal region of the complexed serpin.[157,405]

In addition to examples of signaling properties of serpin−proteinase complexes mediated by LDL receptor family receptors, there are a number of examples of uncomplexed serpins having receptor-mediated biological activities, presumably by some other mechanism. Examples of this are the antiangiogenic properties associated with cleaved and latent conformations of antithombin,[156] and the antiangiogenic[406,407] and neurotrophic properties of PEDF.[408] At least in the case of PEDF, there is evidence for the involvement of a specific receptor.[409]

Finally, there are examples of serpins serving as sources of proteolytically derived peptides. Leaving aside angiotensinogen, for which renin cleavage to

generate the prohormone decapeptide angiotensin I appears to be the principal raison d'etre of the serpin, there is the example of generation of bioactive peptides from heparin cofactor II.[410,411]

7. Concluding Remarks

At the level of protein biosynthesis and folding, the serpin fold represents a fine balance between conflicting needs of, on one hand, features that are needed to direct metastable folding and the ability to undergo facile conformational interconversion as part of the normal mechanism of proteinase inhibition, and on the other hand, the avoidance of inappropriate or premature inactivation through conversion to the latent state or through polymerization. That this is a fine balance is illustrated by the many spontaneous single point mutations in many serpins that have been identified through the adverse effects they have on functional properties and, in humans, the many instances of pathologies associated with these mutations. Despite this, serpins are extremely widely distributed and very abundant, and appear frequently to be the proteinase inhibitors of choice for regulation of proteinase-dependent pathways in multicellular organisms. This indicates that there must be significant advantages of the serpins over lock-and-key type inhibitors for them to be maintained, and in many instances multiplied by gene duplication, over many millions of years.

The mechanism of proteinase inhibition, being dependent on a mechanical mechanism rather than a more normal specificity-of-interaction mechanism, gives clear advantages in allowing inhibition of many serine proteinases, even of different tertiary folds, and also of cysteine proteinases of very different folds. At the same time, the importance of the initial recognition of proteinase in determining the rate of inhibition allows for regulation of these rates by use of both the reactive center loop and exosite interactions, while modulation of the relative fluxes along different branches of the reaction pathway allows modulation of outcome. In addition, the conformational changes that accompany reaction allow for opportunities of signaling, whether the serpin is an inhibitory or noninhibitory one.

In the future, studies on serpins are likely to identify many more species that interact with them and influence the rate and outcome of a reaction with proteinase, the localization of such interactions, and the accessibility of the serpin to its targets. For many serpins, the true targets are unknown and will need to be identified if the true in vivo roles of these serpins are to be fully elucidated.

8. Abbreviations

α_1PI	α_1-proteinase inhibitor
FRET	fluorescence resonance energy transfer
GAG	glycosaminoglycan
HCII	heparin cofactor II
MMP	matrix metalloproteinase
PAI-1	plasminogen activator inhibitor-1
P1−P1′etc.	designation of residues in the reactive center loop, using the nomenclature of Schechter and Berger,[412] in which the scissile bond is between residues P1 and P1′, residues N-terminal to this are designated P2, P3, etc., and those C-terminal P2′, P3′, etc.
PEDF	pigment epithelium derived factor
RAP	receptor associated protein
RCL	reactive center loop
scuPA	single-chain urokinase-type plasminogen activator
tcuPA	two-chain urokinase-type plasminogen activator
tPA	tissue-type plasminogen activator
uPAR	urokinase plasminogen activator receptor
ZPI	protein Z-dependent proteinase inhibitor
LRP	low-density lipoprotein receptor associated protein

9. Acknowledgments

I thank Jim Huntington for providing data on the structure of heparin cofactor II prior to publication, James Whisstock for providing information on the existence of bacterial serpins prior to publication, Miljan Simonovic for preparing the structural figures, and Steven Olson for helpful comments on the manuscript. I thank the National Institutes of Health for support through Grants HL49234, HL64013, and GM54414.

10. References

(1) Hunt, L. T.; Dayhoff, M. O. *Biochem. Biophys. Res. Commun.* **1980**, *95*, 864.
(2) Wright, H. T. *J. Biol. Chem.* **1984**, *259*, 14335.
(3) Bock, S. C.; Skriver, K.; Nielsen, E.; Thøgersen, H. C.; Wiman, B.; Donaldson, V. H.; Eddy, R. L.; Marrinan, J.; Radziejewska, E.; Huber, R.; Shows, T. B.; Magnusson, S. *Biochemistry* **1986**, *25*, 4292.
(4) Irving, J. A.; Pike, R. N.; Lesk, A. M.; Whisstock, J. C. *Genome Res.* **2000**, *10*, 1845.
(5) Carrell, R. W.; Travis, J. *Trends Biol. Sci.* **1985**, *10*, 20.
(6) Doolittle, R. F. *Science* **1983**, *222*, 417.
(7) Silverman, G. A.; Bird, P. I.; Carrell, R. W.; Church, F. C.; Coughlin, P. B.; Gettins, P. G. W.; Irving, J. A.; Lomas, D. A.; Luke, C. J.; Moyer, R. W.; Pemberton, P. A.; Remold-O'Donnell, E.; Salvesen, G. S.; Travis, J.; Whisstock, J. C. *J. Biol. Chem.* **2001**, *276*, 33293.
(8) Remold-O'Donnell, E. *FEBS Lett.* **1993**, *315*, 105.
(9) Scott, F. L.; Eyre, H. J.; Lioumi, M.; Ragoussis, J.; Irving, J. A.; Sutherland, G. A.; Bird, P. I. *Genomics* **1999**, *62*, 490.
(10) Abts, H. F.; Welss, T.; Scheuring, S.; Scott, F. L.; Irving, J. A.; Michel, G.; Bird, P. I.; Ruzicka, T. *DNA Cell Biol.* **2001**, *20*, 123.
(11) Spring, P.; Nakashima, T.; Frederick, M.; Henderson, Y.; Clayman, G. *Biochem. Biophys. Res. Commun.* **1999**, *264*, 299.
(12) Zou, Z.; Anisowicz, A.; Hendrix, M. J. C.; Thor, A.; Neveu, M.; Sheng, S.; Rafidi, K.; Seftor, E.; Sager, R. *Science* **1994**, *263*, 526.
(13) Sheng, S.; Carey, J.; Seftor, E. A.; Dias, L.; Hendrix, M. J. C.; Sager, R. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 11669.
(14) Blacque, O. E.; Worrall, D. M. *J. Biol. Chem.* **2002**, *277*, 10783.
(15) Schneider, S. S.; Schick, C.; Fish, K. E.; Miller, E.; Pena, J. C.; Treter, S. D.; Hui, S. M.; Silverman, G. A. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 3147.
(16) Jiang, H. B.; Wang, Y.; Huang, Y. L.; Mulnix, A. B.; Kadel, J.; Cole, K.; Kanost, M. R. *J. Biol. Chem.* **1996**, *271*, 28017.
(17) Irving, J. A.; Steenbakkers, P. J. M.; Lesk, A. M.; Op den Camp, H. J.; Pike, R. N.; Whisstock, J. C. **2002**, submitted.
(18) Belin, D.; Wohlwend, A.; Schleunung, W.-D.; Kruithof, E. K. O.; Vassalli, J.-D. *EMBO J.* **1989**, *8*, 3287.
(19) Mikus, P.; Urano, T.; Liljeström, P.; Ny, T. *Eur. J. Biochem.* **1993**, *218*, 1071.
(20) Grigoryev, S. A.; Bednar, J.; Woodcock, C. L. *J. Biol. Chem.* **1999**, *274*, 5626.
(21) Deleted in proof.
(22) Stein, P. E.; Carrell, R. W. *Nat. Struct. Biol.* **1995**, *2*, 96.

(23) Jakubowski, H. V.; Kline, M. D.; Owen, W. G. *J. Biol. Chem.* **1986**, *261*, 3876.
(24) Ishiguro, K.; Kojima, T.; Kadomatsu, K.; Nakayama, Y.; Takagi, A.; Suzuki, M.; Takeda, N.; Ito, M.; Yamamoto, K.; Matsushita, T.; Kusugami, K.; Muramatsu, T.; Saito, H. *J. Clin. Invest.* **2000**, *106*, 873.
(25) Cox, D. W. In *The Metabolic Basis of Inherited Disease*; 6th ed.; Scriver, C. R., Beaudet, A. L., Sly, W. S., Valle, D., Eds.; McGraw-Hill: New York, 1989.
(26) Beatty, K.; Bieth, J.; Travis, J. *J. Biol. Chem.* **1980**, *255*, 3931.
(27) Wiman, B.; Collen, D. *Eur. J. Biochem.* **1978**, *84*, 573.
(28) Lijnen, H. R.; Okada, K.; Matsuo, O.; Collen, D.; Dewerchin, M. *Blood* **1999**, *93*, 2274.
(29) Davis, R. L.; Shrimpton, A. E.; Holohan, P. D.; Bradshaw, C.; Feiglin, D.; Collins, G. H.; Sonderegger, P.; Kinter, J.; Becker, L. M.; Lacbawan, F.; Krasnewich, D.; Muenke, M.; Lawrence, D. A.; Yerby, M. S.; Shaw, C. M.; Gooptu, B.; Elliott, P. R.; Finch, J. T.; Carrell, R. W.; Lomas, D. A. *Nature (London)* **1999**, *401*, 376.
(30) Krueger, S. R.; Ghisu, G. P.; Cinelli, P.; Gschwend, T. P.; Osterwalder, T.; Wolfer, D. P.; Sonderegger, P. *J. Neurosci.* **1997**, *17*, 8984.
(31) Pemberton, P. A.; Stein, P. E.; Pepys, M. B.; Potter, J. M.; Carrell, R. W. *Nature (London)* **1988**, *336*, 257.
(32) Suda, S. A.; Gettins, P. G. W.; Patston, P. A. *Arch. Biochem. Biophys.* **2000**, *384*, 31.
(33) Nagai, N.; Hosokawa, M.; Itohara, S.; Adachi, E.; Matsushita, T.; Hosokawa, N.; Nagata, K. *J. Cell Biol.* **2000**, *150*, 1499.
(34) Tempfer, C. B.; Moreno, R. M.; Gregg, A. R. *Biol. Reprod.* **2000**, *62*, 457.
(35) Travis, J.; Salvesen, G. *Annu. Rev. Biochem.* **1983**, *52*, 655.
(36) Inglis, J. D.; Hill, R. E. *EMBO J.* **1991**, *10*, 255.
(37) Copeland, N. G.; Jenkins, N. A.; Gilbert, D. J.; Eppig, J. T.; Maltais, L. J.; Miller, J. C.; Dietrich, W. F.; Weaver, A.; Lincoln, S. E.; Steen, R. G.; Stein, L. D.; Nadeau, J. H.; Lander, E. S. *Science* **1993**, *262*, 57.
(38) Sun, J.; Rose, J. B.; Bird, P. *J. Biol. Chem.* **1995**, *270*, 16089.
(39) Sun, J. R.; Ooms, L.; Bird, C. H.; Sutton, V. R.; Trapani, J. A.; Bird, P. I. *J. Biol. Chem.* **1997**, *272*, 15434.
(40) Kaiserman, D.; Knaggs, S.; Scarff, K. L.; Gillard, A.; Mirza, G.; Cadman, M.; McKeone, R.; Denny, P.; Cooley, J.; Benarafa, C.; Remold-O'Donnell, E.; Ragoussis, J.; Bird, P. I. *Genomics* **2002**, *79*, 349.
(41) Zhang, J.; Broze, G. J., Jr. *Thromb. Haemostasis* **2001**, *85*, 861.
(42) Wu, J. K.; Sheffield, W. P.; Blajchman, M. A. *Thromb. Haemostasis* **1992**, *68*, 291.
(43) Zhang, G.-S.; Mehringer, J. H.; Van Deerlin, V. M. D.; Kozak, C. A.; Tollefsen, D. M. *Biochemistry* **1994**, *33*, 3632.
(44) Carmeliet, P.; Kieckens, L.; Schoonjans, L.; Ream, B.; van Nuffelen, A.; Prendergast, G.; Cole, M.; Bronson, R.; Collen, D.; Mulligan, R. C. *J. Clin. Invest.* **1993**, *92*, 2746.
(45) Carter, R. E.; Cerosaletti, K. M.; Burkin, D. J.; Fournier, R. E. K.; Jones, C.; Greenberg, B. D.; Citron, B. A.; Festoff, B. W. *Genomics* **1995**, *27*, 196.
(46) Kozaki, K.; Miyaishi, O.; Koiwai, O.; Yasui, Y.; Kashiwai, A.; Nishikawa, Y.; Shimizu, S.; Saga, S. *J. Biol. Chem.* **1998**, *273*, 15125.
(47) Shirozu, M.; Tada, H.; Tashiro, K.; Nakamura, T.; Lopez, N. D.; Nazarea, M.; Hamada, T.; Sato, T.; Nakano, T.; Honjo, T. *Genomics* **1996**, *37*, 273.
(48) Lener, M.; Vinci, G.; Duponchel, C.; Meo, T.; Tosi, M. *Eur. J. Biochem.* **1998**, *254*, 117.
(49) Dafforn, T. R.; Della, M.; Miller, A. D. *J. Biol. Chem.* **2001**, *276*, 49310.
(50) Schwartzenberg, S. J.; Yoon, J. B.; Sharp, H. L.; Seelig, S. *Am. J. Physiol.* **1989**, *256*, C413.
(51) Ostergaard, H.; Rasmussen, S. K.; Roberts, T. H.; Hejgaard, J. *J. Biol. Chem.* **2000**, *275*, 33272.
(52) Hejgaard, J. *FEBS Lett.* **2001**, *488*, 149.
(53) Dahl, S. W.; Rasmussen, S. K.; Hejgaard, J. *J. Biol. Chem.* **1996**, *271*, 25083.
(54) Dahl, S. W.; Rasmussen, S. K.; Petersen, L. C.; Hejgaard, J. *FEBS Lett.* **1996**, *394*, 165.
(55) Yoo, B.-C.; Aoki, K.; Xiang, Y.; Campbell, L. R.; Hull, R. J.; Xoconostle-Cazares, B.; Monzer, J.; Lee, J.-Y.; Ullman, D. E.; Lucas, W. J. *J. Biol. Chem.* **2000**, *275*, 35122.
(56) Guerin, J. L.; Gelfi, J.; Camus, C.; Delverdier, M.; Whisstock, J. C.; Amardeihl, M. F.; Py, R.; Bertagnoli, S.; Messud-Petit, F. *J. Gen. Virol.* **2001**, *82*, 1407.
(57) Turner, P. C.; Baquero, M. T.; Yuan, S.; Thoennes, S. R.; Moyer, R. W. *Virology* **2000**, *272*, 267.
(58) Afonso, C. L.; Tulman, E. R.; Lu, Z.; Zsak, L.; Osorio, F. A.; Balinsky, C.; Kutish, G. F.; Rock, D. L. *J. Virol.* **2002**, *76*, 783.
(59) Coleman, S.; Drahn, B.; Petersen, G.; Stolorov, J.; Kraus, K. *Insect Biochem. Mol. Biol.* **1995**, *25*, 203.
(60) Okuyama, E.; Tachida, H.; Yamazaki, T. *J. Mol. Evolution* **1997**, *45*, 32.
(61) Misra, S.; Hecht, P.; Maeda, R.; Anderson, K. V. *Development* **1998**, *125*, 1261.
(62) Wolfner, M. F.; Harada, H. A.; Bertram, M. J.; Stelick, T. J.; Kraus, K. W.; Kalb, J. M.; Lung, Y. O.; Neubaum, D. M.; Park, M.; Tram, U. *Insect Biochem. Mol. Biol.* **1997**, *27*, 825.
(63) Levashina, E. A.; Langley, E.; Green, C.; Gubb, D.; Ashburner, M.; Hoffman, J. A.; Reichhart, J.-M. *Science* **1999**, *285*, 1917.
(64) Green, C.; Levashina, E. A.; McKinnie, C.; Dafforn, T. R.; Reichhart, J.-M.; Gubb, D. *Genetics* **2000**, *156*, 1117.
(65) Jiang, H. B.; Kanost, M. R. *J. Biol. Chem.* **1997**, *272*, 1082.
(66) Sasaki, T. *Eur. J. Biochem.* **1991**, *202*, 255.
(67) Narumi, H.; Hishida, T.; Sasaki, T.; Feng, D.-F.; Doolittle, R. F. *Eur. J. Biochem.* **1993**, *214*, 181.
(68) Agarwala, K. L.; Kawabata, S.; Miura, Y.; Kuroki, Y.; Iwanaga, S. *J. Biol. Chem.* **1996**, *271*, 23768.
(69) Stark, K. R.; James, A. A. *J. Biol. Chem.* **1998**, *273*, 20802.
(70) Whisstock, J. C.; Irving, J. A.; Bottomley, S. P.; Pike, R. N.; Lesk, A. M. *Proteins* **1999**, *36*, 31.
(71) Blanton, R. E.; Licate, L. S.; Aman, R. A. *Mol. Biochem. Parasitol.* **1994**, *63*, 1.
(72) Li, Z.; King, C. L.; Ogundipe, J. O.; Licate, L. S.; Blanton, R. E. *J. Inf. Dis.* **1995**, *171*, 416.
(73) Yenbutr, P.; Scott, A. L. *Infect. Immunol.* **1995**, *63*, 1745.
(74) Colwell, N. S.; Tollefsen, D. M. *Thromb. Haemostasis* **1998**, *80*, 784.
(75) Hirayoshi, K.; Kudo, H.; Takechi, H.; Nakai, A.; Iwamatsu, A.; Yamada, K.; Nagata, K. *Mol. Cell. Biol.* **1991**, *11*, 4036.
(76) Osterwalder, T.; Contartese, J.; Stoeckli, E. T.; Kuhn, T. B.; Sonderegger, P. *EMBO J.* **1996**, *15*, 2944.
(77) Backovic, M.; Gettins, P. G. W. *J. Proteome Res.* **2002**, *1*, 367.
(78) Aranishi, F. *Marine Biotechnol.* **1999**, *1*, 81.
(79) Aranishi, F. *Marine Biotechnol.* **1999**, *1*, 33.
(80) Huang, C.-J.; Chen, C.-C.; Chen, H.-J.; Huang, F.-L.; Chang, G.-D. *J. Neurochem.* **1995**, *64*, 1715.
(81) Huang, C.-J.; Lee, M.-S.; Huang, F.-L.; Chang, G.-D. *J. Neurochem.* **1995**, *64*, 1721.
(82) Andersen, O.; Flengsrud, R.; Norberg, K.; Salte, R. *Eur. J. Biochem.* **2000**, *267*, 1651.
(83) Holland, L. J.; Suksang, C.; Wall, A. A.; Roberts, L. R.; Moser, D. R.; Bhattacharya, A. *J. Biol. Chem.* **1992**, *267*, 7053.
(84) Löbermann, H.; Tokuoka, R.; Deisenhofer, J.; Huber, R. *J. Mol. Biol.* **1984**, *177*, 731.
(85) Stein, P. E.; Leslie, A. G. W.; Finch, J. T.; Turnell, W. G.; McLaughlin, P. J.; Carrell, R. W. *Nature (London)* **1990**, *347*, 99.
(86) Wright, H. T.; Qian, H. X.; Huber, R. *J. Mol. Biol.* **1990**, *213*, 513.
(87) Mottonen, J.; Strand, A.; Symersky, J.; Sweet, R. M.; Danley, D. E.; Georghegan, K. F.; Gerard, R. D.; Goldsmith, E. J. *Nature (London)* **1992**, *355*, 270.
(88) Schreuder, H. A.; de Boer, B.; Dijkema, R.; Mulders, J.; Theunissen, H. J. M.; Grootenhuis, P. D. J.; Hol, W. G. J. *Nat. Struct. Biol.* **1994**, *1*, 48.
(89) Carrell, R. W.; Stein, P. E.; Fermi, G.; Wardell, M. R. *Structure* **1994**, *2*, 257.
(90) Wei, A.; Rubin, H.; Cooperman, B. S.; Christianson, D. W. *Natu. Struct. Biol.* **1994**, *1*, 251.
(91) Harrop, S. J.; Jankova, L.; Coles, M.; Jardine, D.; Whittaker, J. S.; Gould, A. R.; Meister, A.; King, G. C.; Mabbutt, B. C.; Curmi, P. M. G. *Structure* **1999**, *7*, 43.
(92) Simonovic, M.; Gettins, P. G. W.; Volz, K. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 11131.
(93) Renatus, M.; Zhou, Q.; Stennike, H. R.; Snipas, S. J.; Turk, D.; Bankston, L. A.; Liddington, R. C.; Salvesen, G. S. *Structure* **2000**, *8*, 789.
(94) Simonovic, M.; Gettins, P. G. W.; Volz, K. *Protein Sci.* **2000**, *9*, 1423.
(95) Yamasaki, M.; Arii, Y.; Mikami, B.; Hirose, M. *J. Mol. Biol.* **2002**, *315*, 113.
(96) Li, J.; Wang, Z.; Canagarajah, B.; Jiang, H.; Kanost, M.; Goldsmith, E. J. *Structure* **1999**, *7*, 103.
(97) Elliott, P. R.; Lomas, D. A.; Carrell, R. W.; Abrahams, J. P. *Nat. Struct. Biol.* **1996**, *3*, 676.
(98) Elliott, P. R.; Abrahams, J. P.; Lomas, D. A. *J. Mol. Biol.* **1998**, *275*, 419.
(99) Kim, S.-J.; Woo, J.-R.; Seo, E. J.; Yu, M.-H.; Ryu, S.-E. *J. Mol. Biol.* **2001**, *306*, 109.
(100) Huntington, J. A.; McCoy, A.; Pei, X. Y.; Gettins, P. G. W.; Carrell, R. W. *J. Biol. Chem.* **2000**, *275*, 15377.
(101) Skinner, R.; Abrahams, J.-P.; Whisstock, J. C.; Lesk, A. M.; Carrell, R. W.; Wardell, M. R. *J. Mol. Biol.* **1997**, *266*, 601.
(102) Danielsson, Å.; Björk, I. *Biochem. J.* **1982**, *207*, 21.
(103) Huntington, J. A.; Olson, S. T.; Fan, B.; Gettins, P. G. W. *Biochemistry* **1996**, *35*, 8495.
(104) Baglin, T. P.; Carrell, R. W.; Church, F. C.; Esmon, C. T.; Huntington, J. A. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 11079.
(105) Chuang, T. L.; Schleef, R. R. *J. Biol. Chem.* **1999**, *274*, 11194.
(106) Dickinson, J. L.; Norris, B. J.; Jensen, P. H.; Antalis, T. M. *Cell Death Differ.* **1998**, *5*, 163.

(107) Jensen, P. H.; Schüler, E.; Woodrow, G.; Richardson, M.; Goss, N.; Hojrup, P.; Petersen, T. E.; Rasmussen, L. K. *J. Biol. Chem.* **1994**, *269*, 15394.
(108) Stein, P. E.; Chothia, C. *J. Mol. Biol.* **1991**, *221*, 615.
(109) Kaslik, G.; Kardos, J.; Szabó, L.; Závodszky, P.; Westler, W. M.; Markley, J. L.; Gráf, L. *Biochemistry* **1997**, *36*, 5455.
(110) Boudier, C.; Bieth, J. G. *Biochemistry* **2001**, *40*, 9962.
(111) Lawrence, D. A.; Olson, S. T.; Palaniappan, S.; Ginsburg, D. *Biochemistry* **1994**, *33*, 3643.
(112) Gooptu, B.; Chang, W.-S. W.; Dafforn, T. R.; Carrell, R. W.; Read, R. J. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 67.
(113) Kim, D.; Yu, M.-H. *Biochem. Biophys. Res. Commun.* **1996**, *226*, 378.
(114) Koloczek, H.; Banbula, A.; Salvesen, G. S.; Potempa, J. *Protein Sci.* **1996**, *5*, 2226.
(115) Wang, Z. L.; Mottonen, J.; Goldsmith, E. J. *Biochemistry* **1996**, *35*, 16443.
(116) Pearce, M. C.; Rubin, H.; Bottomley, S. P. *J. Biol. Chem.* **2000**, *275*, 28513.
(117) Patston, P. A.; Gettins, P. G. W. *FEBS Lett.* **1996**, *383*, 87.
(118) Lomas, D. A.; Elliott, P. R.; Chang, W.-S. W.; Wardell, M. R.; Carrell, R. W. *J. Biol. Chem.* **1995**, *270*, 5282.
(119) Im, H.; Ahn, H. Y.; Yu, M. H. *Protein Sci.* **2000**, *9*, 1497.
(120) Kwon, K.-S.; Kim, J.; Shin, H.-S.; Yu, M.-H. *J. Biol. Chem.* **1994**, *269*, 9627.
(121) Lee, K. N.; Park, S. D.; Yu, M. H. *Nat. Struct. Biol.* **1996**, *3*, 497.
(122) Nyung, K.; Im, H.; Kang, S. W.; Yu, M.-H. *J. Biol. Chem.* **1998**, *273*, 2509.
(123) Berkenpas, M. B.; Lawrence, D. A.; Ginsburg, D. *EMBO J.* **1995**, *14*, 2969.
(124) Patston, P. A.; Hauert, J.; Michaud, M.; Schapira, M. *FEBS Lett.* **1995**, *368*, 401.
(125) Lomas, D. A.; Evans, D. L.; Finch, J. T.; Carrell, R. W. *Nature (London)* **1992**, *357*, 605.
(126) Elliott, P. R.; Bilton, D.; Lomas, D. A. *Am. J. Respir. Cell Mol. Biol.* **1998**, *18*, 670.
(127) Mahadeva, R.; Lomas, D. A. *Thorax* **1998**, *53*, 501.
(128) Wallaert, B.; Gressier, B.; Marquette, C. H.; Gosset, P.; Remy-jardin, M.; Mizon, J.; Tonnel, A. B. *Am. Rev. Respir. Dis.* **1993**, *147*, 1537.
(129) Ogushi, F.; Hubbard, R. C.; Vogelmeier, C.; Fells, G. A.; Crystal, R. G. *J. Clin. Invest.* **1991**, *87*, 1060.
(130) Crystal, R. G. *J. Clin. Invest.* **1990**, *85*, 1343.
(131) Perlmutter, D. H.; Pierce, J. A. *Am. J. Physiol.* **1989**, *257*, L147.
(132) Sivasothy, P.; Dafforn, T. R.; Gettins, P. G. W.; Lomas, D. A. *J. Biol. Chem.* **2000**, *275*, 33663.
(133) Yu, M.-H.; Lee, K. N.; Kim, J. *Nat. Struct. Biol.* **1995**, *2*, 363.
(134) Bruce, D.; Perry, D. J.; Borg, J.-Y.; Carrell, R. W.; Wardell, M. R. *J. Clin. Invest.* **1994**, *94*, 2265.
(135) Mast, A. E.; Enghild, J. J.; Salvesen, G. S. *Biochemistry* **1992**, *31*, 2720.
(136) Huntington, J. A.; Pannu, N. S.; Hazes, B.; Read, R. J. *J. Mol. Biol.* **1999**, *293*, 449.
(137) Dunstone, M. A.; Dai, W.; Whisstock, J.; Rossjohn, J.; Pike, R. N.; Feil, S. C.; Le Bonniec, B. F.; Parker, M. W.; Bottomley, S. P. *Protein Sci.* **2000**, *9*, 417.
(138) Bode, W.; Huber, R. *Eur. J. Biochem.* **1992**, *204*, 433.
(139) Huntington, J. A.; Read, R. J.; Carrell, R. W. *Nature (London)* **2000**, *407*, 923.
(140) Stratikos, E.; Gettins, P. G. W. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 453.
(141) Mellet, P.; Boudier, C.; Mely, Y.; Bieth, J. G. *J. Biol. Chem.* **1998**, *273*, 9119.
(142) Stratikos, E.; Gettins, P. G. W. *J. Biol. Chem.* **1998**, *273*, 15582.
(143) Stratikos, E.; Gettins, P. G. W. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 4808.
(144) Fa, M.; Bergstrom, F.; Hagglof, P.; Wilczynska, M.; Johansson, L.; Ny, T. *Structure* **2000**, *8*, 397.
(145) Peterson, F. C.; Gettins, P. G. W. *Biochemistry* **2001**, *40*, 6284.
(146) Backovic, M.; Stratikos, E.; Lawrence, D. A.; Gettins, P. G. W. *Protein Sci.* **2002**, *11*, 1182.
(147) Read, R. J.; James, M. N. G. In *Proteinase Inhibitors*; Barrett, A. J., Salvesen, G. S., Eds.; Elsevier Science Publishers: Amsterdam, 1986.
(148) Hubbard, S. J.; Eisenmenger, F.; Thornton, J. M. *Protein Sci.* **1994**, *3*, 757.
(149) Ako, H.; Foster, R. J.; Ryan, C. A. *Biochem. Biophys. Res. Commun.* **1972**, *47*, 1402.
(150) Wilmouth, R. C.; Edman, K.; Neutze, R.; Wright, P. A.; Clifton, I. J.; Schneider, T. R.; Schofield, C. J.; Hajdu, J. *Nat. Struct. Biol.* **2001**, *8*, 689.
(151) Calugaru, S. V.; Swanson, R.; Olson, S. T. *J. Biol. Chem.* **2001**, *276*, 32446.
(152) Hood, D. B.; Huntington, J. A.; Gettins, P. G. W. *Biochemistry* **1994**, *33*, 8538.
(153) Patston, P. A.; Gettins, P.; Beechem, J.; Schapira, M. *Biochemistry* **1991**, *30*, 8876.
(154) Hermans, J. M.; Monard, D.; Jones, R.; Stone, S. R. *Biochemistry* **1995**, *34*, 3678.
(155) Olson, S. T. *J. Biol. Chem.* **1985**, *260*, 10153.
(156) O'Reilly, M. S.; Pirie-Shepherd, S.; Lane, W. S.; Folkman, J. *Science* **1999**, *285*, 1926.
(157) Banda, M. J.; Rice, A. G.; Griffin, G. L.; Senior, R. M. *J. Biol. Chem.* **1988**, *263*, 4481.
(158) Potempa, J.; Fedak, D.; Dubin, A.; Travis, J. *J. Biol. Chem.* **1991**, *266*, 21482.
(159) Peterson, F. C.; Gordon, N. C.; Gettins, P. G. W. *Biochemistry* **2000**, *39*, 11884.
(160) Ye, S.; Cech, A. L.; Belmares, R.; Bergstrom, R. C.; Tong, Y.; Corey, D. R.; Kanost, M.; Goldsmith, E. J. *Nat. Struct. Biol.* **2001**, *8*, 979.
(161) Bode, W.; Huber, R. *Curr. Opin. Struct. Biol.* **1991**, *1*, 45.
(162) Olson, S. T.; Swanson, R.; Day, D.; Verhamme, I.; Kvassman, J.; Shore, J. D. *Biochemistry* **2001**, *40*, 11742.
(163) Mellet, P.; Bieth, J. G. *J. Biol. Chem.* **2000**, *275*, 10788.
(164) Jesty, J. *J. Biol. Chem.* **1979**, *254*, 10044.
(165) Jörnvall, H.; Fish, W. W.; Björk, I. *FEBS Lett.* **1979**, *106*, 358.
(166) Longas, M. O.; Finlay, T. H. *Biochem. J.* **1980**, *189*, 481.
(167) Fish, W. W.; Björk, I. *Eur. J. Biochem.* **1979**, *101*, 31.
(168) Björk, I.; Jackson, C. M.; Jörnvall, H.; Lavine, K. K.; Nordling, K.; Salsgiver, W. J. *J. Biol. Chem.* **1982**, *257*, 2406.
(169) Matheson, N. R.; van Halbeek, H.; Travis, J. *J. Biol. Chem.* **1991**, *266*, 13489.
(170) Whisstock, J.; Lesk, A. M.; Carrell, R. *Proteins* **1996**, *26*, 288.
(171) Olson, S. T.; Bock, P. E.; Kvassman, J.; Shore, J. D.; Lawrence, D. A.; Ginsburg, D.; Björk, I. *J. Biol. Chem.* **1995**, *270*, 30007.
(172) Olson, S. T.; Swanson, R.; Patston, P. A.; Björk, I. *J. Biol. Chem.* **1997**, *272*, 13338.
(173) Lawrence, D. A.; Ginsburg, D.; Day, D. E.; Berkenpas, M. B.; Verhamme, I. M.; Kvassman, J.-O.; Shore, J. D. *J. Biol. Chem.* **1995**, *270*, 25309.
(174) Wright, H. T.; Scarsdale, J. N. *Proteins* **1995**, *22*, 210.
(175) Gettins, P. G. W.; Patston, P. A.; Olson, S. T. *Serpins: Structure, Function and Biology*; R. G. Landes Co.: Austin, 1996.
(176) Futamura, A.; Stratikos, E.; Olson, S. T.; Gettins, P. G. W. *Biochemistry* **1998**, *37*, 13110.
(177) Plotnick, M. I.; Mayne, L.; Schechter, N. M.; Rubin, H. *Biochemistry* **1996**, *35*, 7586.
(178) Bijnens, A.-P.; Gils, A.; Knockaert, I.; Stassen, J. M.; Declerck, P. J. *J. Biol. Chem.* **2000**, *275*, 6375.
(179) Shore, J. D.; Day, D. E.; Francis-Chmura, A. M.; Verhamme, I.; Kvassmann, J.; Lawrence, D. A.; Ginsburg, D. *J. Biol. Chem.* **1995**, *270*, 5395.
(180) Stratikos, E.; Gettins, P. G. W. *J. Biol. Chem.* **1998**, *273*, 15582.
(181) Kaslik, G.; Patthy, A.; Bálint, M.; Gráf, L. *FEBS Lett.* **1995**, *370*, 179.
(182) Stavridi, E. S.; O'Malley, K.; Lukacs, C. M.; Moore, W. T.; Lambris, J. D.; Christianson, D. W.; Rubin, H.; Cooperman, B. S. *Biochemistry* **1996**, *35*, 10608.
(183) Egelund, R.; Petersen, T. E.; Andreasen, P. A. *Eur. J. Biochem.* **2001**, *268*, 673.
(184) O'Malley, K. M.; Cooperman, B. S. *J. Biol. Chem.* **2001**, *276*, 6631.
(185) Komiyama, T.; Grøn, H.; Pemberton, P. A.; Salvesen, G. S. *Protein Sci.* **1996**, *5*, 874.
(186) Dahlen, J. R.; Foster, D. C.; Kisiel, W. *Biochem. Biophys. Res. Commun.* **1997**, *238*, 329.
(187) Dahlen, J. R.; Jean, F.; Thomas, G.; Foster, D. C.; Kisiel, W. *J. Biol. Chem.* **1998**, *273*, 1851.
(188) Lukacs, C. M.; Zhong, J. Q.; Plotnick, M. I.; Rubin, H.; Cooperman, B. S.; Christianson, D. W. *Nat. Struct. Biol.* **1996**, *3*, 888.
(189) Lawrence, D. A.; Olson, S. T.; Muhammad, S.; Day, D. E.; Kvassman, J.-O.; Ginsburg, D.; Shore, J. D. *J. Biol. Chem.* **2000**, *275*, 5839.
(190) Huntington, J. A.; Patston, P. A.; Gettins, P. G. W. *Protein Sci.* **1995**, *4*, 613.
(191) Huntington, J. A.; Fan, B.; Karlsson, K. E.; Deinum, J.; Lawrence, D. A.; Gettins, P. G. W. *Biochemistry* **1997**, *36*, 5432.
(192) Hopkins, P. C. R.; Carrell, R. W.; Stone, S. R. *Biochemistry* **1993**, *32*, 7650.
(193) Skinner, R.; Chang, W. S. W.; Jin, L.; Pei, X.; Huntington, J. A.; Abrahams, J. P.; Carrell, R. W.; Lomas, D. A. *J. Mol. Biol.* **1998**, *283*, 9.
(194) Xue, Y. F.; Björquist, P.; Inghardt, T.; Linschoten, M.; Musil, D.; Sjölin, L.; Deinum, J. *Structure* **1998**, *6*, 627.
(195) Jankova, L.; Harrop, S. J.; Saunders: D. N.; Andrews, J. L.; Bertram, K. C.; Gould, A. R.; Baker, M. S.; Curmi, P. M. G. *J. Biol. Chem.* **2001**, *276*, 43374.
(196) Björk, I.; Ylinenjärvi, K.; Olson, S. T.; Bock, P. E. *J. Biol. Chem.* **1992**, *267*, 1976.
(197) Björk, I.; Nordling, K.; Larsson, I.; Olson, S. T. *J. Biol. Chem.* **1992**, *267*, 19047.
(198) Schulze, A. J.; Frohnert, P. W.; Engh, R. A.; Huber, R. *Biochemistry* **1992**, *31*, 7560.

(199) Saunders, D. N.; Jankova, L.; Harrop, S. J.; Curmi, P. M. G.; Gould, A. R.; Ranson, M.; Baker, M. S. *J. Biol. Chem.* **2001**, *276*, 43383.
(200) Hopkins, P. C. R.; Whisstock, J. *Science* **1994**, *265*, 1893.
(201) Hopkins, P. C. R.; Stone, S. R. *Biochemistry* **1995**, *34*, 15872.
(202) Tucker, H. M.; Mottonen, J.; Goldsmith, E. J.; Gerard, R. D. *Nat. Struct. Biol.* **1995**, *2*, 442.
(203) Devraj-Kizuk, R.; Chui, D. H. K.; Prochownik, E. V.; Carter, C. J.; Ofosu, F. A.; Blajchman, M. A. *Blood* **1988**, *72*, 1518.
(204) Perry, D. J.; Daly, M.; Harper, P. L.; Tait, R. C.; Price, J.; Walker, I. D.; Carrell, R. W. *FEBS Lett.* **1991**, *285*, 248.
(205) Holmes, W. E.; Lijnen, H. R.; Nelles, L.; Kluft, C.; Niewenhuis, H. K.; Rijken, D. C.; Collen, D. *Science* **1987**, *238*, 209.
(206) Kluft, C.; Nieuwenhuis, H. K.; Rijken, D. C.; Groonenveld, E.; Wijngaards, G.; van Berkel, W.; Dooijewaard, G.; Sixma, J. J. *J. Clin. Invest.* **1987**, *80*, 1391.
(207) Rijken, D. C.; Groeneveld, E.; Kluft, C.; Nieuwenhuis, H. K. *Biochem. J.* **1988**, *255*, 609.
(208) Zhou, A.; Carrell, R. W.; Huntington, J. A. *J. Biol. Chem.* **2001**, *276*, 27541.
(209) Gettins, P. G. W. *FEBS Lett.* **2002**, *523*, 2.
(210) Hervé, M.; Ghélis, C. *Arch. Biochem. Biophys.* **1991**, *285*, 142.
(211) Picard, V.; Marque, P.-E.; Paolucci, F.; Aiach, M.; Le Bonniec, B. F. *J. Biol. Chem.* **1999**, *274*, 4586.
(212) Plotnick, M. I.; Samakur, M.; M., W. Z.; Liu, X.; Rubin, H.; Schechter, N. M.; Selwood, T. *Biochemistry* **2002**, *41*, 334.
(213) Oda, K.; Laskowski, M.; Kress, L. F.; Kowalski, D. *Biochem. Biophys. Res. Commun.* **1977**, *76*, 1062.
(214) Fish, W. W.; Orre, K.; Björk, I. *Eur. J. Biochem.* **1979**, *101*, 39.
(215) Cooperman, B. S.; Stavridi, E.; Nickbarg, E.; Rescorla, E.; Schechter, N. M.; Rubin, H. *J. Biol. Chem.* **1993**, *268*, 23616.
(216) Björk, I.; Nordling, K.; Raub-Segall, E.; Hellman, U.; Olson, S. T. *Biochem. J.* **1998**, *335*, 701.
(217) Travis, J.; Matheson, N. R.; George, P. M.; Carrell, R. W. *Biol. Chem. Hoppe-Seyler* **1986**, *367*, 853.
(218) Owen, M. C.; Brennan, S. O.; Lewis, J. H.; Carrell, R. W. *N. Engl. J. Med.* **1983**, *309*, 694.
(219) Jallat, S.; Carvallo, D.; Tessier, L. H.; Roecklin, D.; Roitsch, C.; Ogushi, F.; Crystal, R. G.; Courtney, M. *Protein Eng.* **1986**, *1*, 29.
(220) Djie, M. Z.; Le Bonniec, B. F.; Hopkins, P. C. R.; Hipler, K.; Stone, S. R. *Biochemistry* **1996**, *35*, 11461.
(221) Chen, V. C.; Chao, L.; Chao, J. *J. Biol. Chem.* **2000**, *275*, 38457.
(222) Le Bonniec, B. F.; Myles, T.; Johnson, T.; Knight, C. G.; Tapparelli, C.; Stone, S. R. *Biochemistry* **1996**, *35*, 7114.
(223) Dufour, E. K.; Denault, J.-B.; Bissonnette, L.; Hopkins, P. C. R.; Lavigne, P.; Leduc, R. *J. Biol. Chem.* **2001**, *276*, 38971.
(224) Anderson, E. D.; Thomas, L.; Hayflick, J. S.; Thomas, G. *J. Biol. Chem.* **1993**, *268*, 24887.
(225) Plotnick, M. I.; Schechter, N. M.; Wang, Z. M.; Liu, X. Z.; Rubin, H. *Biochemistry* **1997**, *36*, 14601.
(226) McRae, B.; Nakajima, K.; Travis, J.; Powers, J. C. *Biochemistry* **1980**, *19*, 3973.
(227) Olson, S. T.; Björk, I.; Sheffer, R.; Craig, P. A.; Shore, J. D.; Choay, J. *J. Biol. Chem.* **1992**, *267*, 12528.
(228) Bedsted, T.; Swanson, R.; Petitou, M.; Björk, I.; Olson, S. T. 2nd International Conference on Protease Inhibitors, Freising, Germany, 2001.
(229) Chuang, Y.-J.; Swanson, R.; Raja, S. M.; Olson, S. T. *J. Biol. Chem.* **2001**, *276*, 14961.
(230) Chuang, Y.-J.; Swanson, R.; Raja, S. M.; Bock, S. C.; Olson, S. T. *Biochemistry* **2001**, *40*, 6670.
(231) Jin, L.; Abrahams, J. P.; Skinner, R.; Petitou, M.; Pike, R. N.; Carrell, R. W. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 14683.
(232) Nordenman, B.; Danielsson, Å.; Björk, I. *Eur. J. Biochem.* **1978**, *90*, 1.
(233) Jordan, R.; Beeler, D.; Rosenberg, R. D. *J. Biol. Chem.* **1979**, *254*, 2902.
(234) Olson, S. T.; Srinivasan, K. R.; Björk, I.; Shore, J. D. *J. Biol. Chem.* **1981**, *256*, 11073.
(235) Lindahl, U.; Bäckström, G.; Thunberg, L.; Leder, I. G. *Proc. Natl. Acad. Sci. U.S.A.* **1980**, *77*, 6551.
(236) Casu, B.; Oreste, P.; Torri, G.; Zoppetti, G.; Choay, J.; Lormeau, J.-C.; Petitou, M.; Sinaÿ, P. *Biochem. J.* **1981**, *197*, 599.
(237) Thunberg, L.; Bäckström, G.; Lindahl, U. *Carbohydr. Res.* **1982**, *100*, 393.
(238) Choay, J.; Petitou, M.; Lormeau, J. C.; Sinaÿ, P.; Casu, B.; Gatti, G. *Biochem. Biophys. Res. Commun.* **1983**, *116*, 492.
(239) Atha, D. H.; Stephens, A. W.; Rosenberg, R. D. *Proc. Natl. Acad. Sci. U.S.A.* **1984**, *81*, 1030.
(240) Kuhn, L. A.; Griffin, J. H.; Fisher, C. L.; Greengard, J. S.; Bouma, B. N.; España, F.; Tainer, J. A. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 8506.
(241) Jordan, R. E.; Oosta, G. M.; Gardner, W. T.; Rosenberg, R. D. *J. Biol. Chem.* **1980**, *255*, 10073.
(242) Olson, S. T.; Björk, I. *J. Biol. Chem.* **1991**, *266*, 6353.
(243) Rezaie, A. R. *J. Biol. Chem.* **1998**, *273*, 16824.
(244) Rezaie, A. R.; Olson, S. T. *Biochemistry* **2000**, *39*, 12083.
(245) Grootenhuis, P. D. J.; Westerduin, P.; Meuleman, D.; Petitou, M.; van Boeckel, C. A. A. *Nat. Struct. Biol.* **1995**, *2*, 736.
(246) Streusand, V. J.; Björk, I.; Gettins, P. G. W.; Petitou, M.; Olson, S. T. *J. Biol. Chem.* **1995**, *270*, 9043.
(247) Patston, P. A.; Schapira, M. *Blood* **1994**, *84*, 1164.
(248) Wallace, A.; Rovelli, G.; Hofsteenge, J.; Stone, S. R. *Biochem. J.* **1989**, *257*, 191.
(249) Tollefsen, D. M.; Petska, C. A.; Monafo, W. J. *J. Biol. Chem.* **1983**, *258*, 6713.
(250) Laurent, T. C.; Tengblad, A.; Thunberg, L.; Höök, M.; Lindahl, U. *Biochem. J.* **1978**, *175*, 691.
(251) Lane, D. A.; Denton, J.; Flynn, A. M.; Thunberg, L.; Lindahl, U. *Biochem. J.* **1984**, *218*, 725.
(252) Tollefsen, D. M.; Sugimori, T.; Maimone, M. M. *Semin. Thromb. Hemostasis* **1990**, *16 Suppl*, 66.
(253) van Boeckel, C. A. A.; Grootenhuis, P. D. J.; Visser, A. *Nat. Struct. Biol.* **1994**, *1*, 423.
(254) Meagher, J. L.; Huntington, J. A.; Fan, B.; Gettins, P. G. W. *J. Biol. Chem.* **1996**, *271*, 29353.
(255) Futamura, A.; Gettins, P. G. W. *J. Biol. Chem.* **2000**, *275*, 4092.
(256) Futamura, A.; Beechem, J. M.; Gettins, P. G. W. *Biochemistry* **2001**, *40*, 6680.
(257) Belzar, K. J.; Zhou, A.; Carrell, R. W.; Gettins, P. G. W.; Huntington, J. A. *J. Biol. Chem.* **2002**, *277*, 8551.
(258) Meagher, J. L.; Olson, S. T.; Gettins, P. G. W. *J. Biol. Chem.* **2000**, *275*, 2698.
(259) Desai, U. R.; Petitou, M.; Björk, I.; Olson, S. T. *J. Biol. Chem.* **1998**, *273*, 7478.
(260) Desai, U. R.; Petitou, M.; Björk, I.; Olson, S. T. *Biochemistry* **1998**, *37*, 13033.
(261) Kridel, S. J.; Knauer, D. J. *J. Biol. Chem.* **1997**, *272*, 7656.
(262) Arocas, V.; Bock, S. C.; Raja, S. M.; Olson, S. T.; Björk, I. *J. Biol. Chem.* **2001**, *276*, 43809.
(263) Schedin-Weiss, S.; Desai, U. R.; Bock, S. C.; Gettins, P. G. W.; Olson, S. T.; Björk, I. *Biochemistry* **2002**, *41*, 4779.
(264) Desai, U. R.; Swanson, R.; Bock, S. C.; Björk, I.; Olson, S. T. *J. Biol. Chem.* **2000**, *275*, 18976.
(265) Arocas, V.; Bock, S. C.; Olson, S. T.; Björk, I. *Biochemistry* **1999**, *38*, 10196.
(266) Arocas, V.; Turk, B.; Bock, S. C.; Olson, S. T.; Björk, I. *Biochemistry* **2000**, *39*, 8512.
(267) Kridel, S. J.; Chan, W. W.; Knauer, D. J. *J. Biol. Chem.* **1996**, *271*, 20935.
(268) Griffith, M. J.; Noyes, C. M.; Tyndall, J. A.; Church, F. C. *Biochemistry* **1985**, *24*, 6777.
(269) Ragg, H. *Nucleic Acids Res.* **1986**, *14*, 1073.
(270) Tollefsen, D. M.; Pestka, C. A.; Monafo, W. J. *J. Biol. Chem.* **1983**, *258*, 6713.
(271) Ragg, H.; Ulshöfer, T.; Gerewitz, J. *J. Biol. Chem.* **1990**, *265*, 22386.
(272) Blinder, M. A.; Andersson, T. R.; Abildgaard, U.; Tollefsen, D. M. *J. Biol. Chem.* **1989**, *264*, 5128.
(273) Blinder, M. A.; Tollefsen, D. M. *J. Biol. Chem.* **1990**, *265*, 286.
(274) Whinna, H. C.; Blinder, M. A.; Szewczyk, M.; Tollefsen, D. M.; Church, F. C. *J. Biol. Chem.* **1991**, *266*, 8129.
(275) Van Deerlin, V. M. D.; Tollefsen, D. M. *J. Biol. Chem.* **1991**, *266*, 20223.
(276) Rydel, T. J.; Ravichandran, K. G.; Tulinsky, A.; Bode, W.; Huber, R.; Roitsch, C.; Fenton, J. W., II. *Science* **1990**, *249*, 277.
(277) Dennis, S.; Wallace, A.; Hofsteenge, J.; Stone, S. R. *Eur. J. Biochem.* **1990**, *188*, 61.
(278) Tollefsen, D. M. *Adv. Exp. Med. Biol.* **1997**, *425*, 35.
(279) Deng, G.; Royle, G.; Wang, S.; Crain, K.; Loskutoff, D. J. *J. Biol. Chem.* **1996**, *271*, 12716.
(280) Lawrence, D. A.; Berkenpas, M. B.; Palaniappan, S.; Ginsburg, D. *J. Biol. Chem.* **1994**, *269*, 15223.
(281) Lawrence, D. A.; Palaniappan, S.; Stefansson, S.; Olson, S. T.; Francis-Chmura, A. M.; Shore, J. D.; Ginsburg, D. *J. Biol. Chem.* **1997**, *272*, 7676.
(282) Lindahl, T. L.; Sigurdardottir, O.; Wiman, B. *Thromb. Haemostasis* **1989**, *62*, 748.
(283) Ehrlich, H. J.; Gebbink, R. K.; Preissner, K. T.; Keijer, J.; Esmon, N. L.; Mertens, K.; Pannekoek, H. *J. Cell Biol.* **1991**, *115*, 1773.
(284) Naski, M. C.; Lawrence, D. A.; Mosher, D. F.; Podor, T. J.; Ginsburg, D. *J. Biol. Chem.* **1993**, *268*, 12367.
(285) Rezaie, A. R. *J. Biol. Chem.* **2001**, *276*, 15567.
(286) Stefansson, S.; Lawrence, D. A. *Nature (London)* **1996**, *383*, 441.
(287) Seiffert, D.; Smith, J. W. *J. Biol. Chem.* **1997**, *272*, 13705.
(288) Han, X.; Fiehler, R.; Broze, G. J., Jr. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 9250.
(289) Han, X.; Huang, Z. F.; Fiehler, R.; Broze, G. J., Jr. *Biochemistry* **1999**, *38*, 11073.
(290) Han, X.; Fiehler, R.; Broze, G. J., Jr. *Blood* **2000**, *96*, 3049.
(291) Sejima, H.; Hayashi, T.; Deyashiki, Y.; Nishioka, J.; Suzuki, K. *Biochem. Biophys. Res. Commun.* **1990**, *171*, 661.
(292) Tabatabai, A.; Fiehler, R.; Broze, G. J., Jr. *Thromb. Haemostasis* **2001**, *85*, 655.
(293) Yin, Z. F.; Huang, Z. F.; Cui, J.; Fiehler, R.; Lasky, N.; Ginsburg, D.; Broze, G. J., Jr. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 6734.

(294) Mast, A. E.; Enghild, J. J.; Nagase, H.; Suzuki, K.; Pizzo, S. V.; Salvesen, G. S. *J. Biol. Chem.* **1991**, *266*, 15810.
(295) Lijnen, H. R.; Arza, B.; Van Hoef, B.; Collen, D.; Declerck, P. J. *J. Biol. Chem.* **2000**, *275*, 37645.
(296) Lijnen, H. R.; Van Hoef, B.; Collen, D. *Biochim. Biophys. Acta* **2001**, *1547*, 206.
(297) Gettins, P.; Harten, B. *Biochemistry* **1988**, *27*, 3634.
(298) Liu, Z.; Zhou, X.; Shapiro, S. D.; Shipley, J. M.; Twining, S. S.; Diaz, L. A.; Senior, R. M.; Werb, Z. *Cell* **2000**, *102*, 647.
(299) Kress, L. F.; Catanese, J.; Hirayama, T. *Biochim. Biophys. Acta* **1984**, *745*, 113.
(300) Janssen, M.; Meier, J.; Freyvogel, T. A. *Toxicon* **1992**, *30*, 985.
(301) Catanese, J. J.; Kress, L. F. *Biochemistry* **1993**, *32*, 509.
(302) Hammond, G. L.; Smith, C. L.; Paterson, N. A. M.; Sibbald, W. J. *J. Clin. Endocrinol. Metab.* **1990**, *71*, 34.
(303) Schick, C.; Pemberton, P. A.; Shi, G. P.; Kamachi, Y.; Cataltepe, S.; Bartuski, A. J.; Gornstein, E. R.; Brömme, D.; Chapman, H. A.; Silverman, G. A. *Biochemistry* **1998**, *37*, 5258.
(304) Schick, C.; Brömme, D.; Bartuski, A. J.; Uemura, Y.; Schechter, N. M.; Silverman, G. A. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 13465.
(305) Hook, V. Y. H.; Purviance, R. T.; Azaryan, A. V.; Hubbard, G.; Krieger, T. J. *J. Biol. Chem.* **1993**, *268*, 20570.
(306) Komiyama, T.; Ray, C. A.; Pickup, D. J.; Howard, A. D.; Thornberry, N. A.; Peterson, E. P.; Salvesen, G. *J. Biol. Chem.* **1994**, *269*, 19331.
(307) Zhou, Q.; Snipas, S.; Orth, K.; Muzio, M.; Dixit, V. M.; Salvesen, G. S. *J. Biol. Chem.* **1997**, *272*, 7797.
(308) Annand, R. R.; Dahlen, J. R.; Sprecher, C. A.; de Dreu, P.; Foster, D. C.; Mankovich, J. A.; Talanian, R. V.; Kisiel, W.; Giegel, D. A. *Biochem. J.* **1999**, *342*, 655.
(309) Tewari, M.; Quan, L. T.; O'Rourke, K.; Desnoyers, S.; Zeng, Z.; Beidler, D. R.; Poirier, G. G.; Salvesen, G. S.; Dixit, V. M. *Cell* **1995**, *81*, 801.
(310) Luke, C. J.; Schick, C.; Tsu, C.; Whisstock, J.; Irving, J. A.; Brömme, D.; Juliano, L.; Shi, G. P.; Chapman, H. A.; Silverman, G. A. *Biochemistry* **2000**, *39*, 7081.
(311) Wilson, K. P.; Black, J. F.; Thomson, J. A.; Kim, E. E.; Griffith, J. P.; Navia, M. A.; Murcko, M. A.; Chambers, S. P.; Aldape, R. A.; Raybuck, S. A.; Livingston, D. J. *Nature (London)* **1994**, *370*, 270.
(312) Lee, D.; Long, S. A.; Adams, J. L.; Chan, G.; Vaidya, K. S.; Francis, T. A.; Kikly, K.; Winkler, J. D.; Sung, C. M.; Debouck, C.; Richardson, S.; Levy, M.; DeWolf, W. E. J.; Keller, P. M.; Tomaszek, T.; Head, M. S.; Ryan, M. D.; Haltiwanger, R. C.; Liang, P. H.; Janson, C. A.; McDevitt, P. J.; Johanson, K.; Concha, N. O.; Chan, W.; Abdel-Meguid, S. S.; Badger, A. M.; Lark, M. W.; Nadeau, D. P.; Suv, L. J.; Gowen, M.; Nuttall, M. E. *J. Biol. Chem.* **2000**, *275*, 16007.
(313) Wei, Y. Y.; Fox, T.; Chambers, S. P.; Sintchak, J.; Coll, J. T.; Golec, J. M. C.; Swenson, L.; Wilson, K. P.; Charifson, P. S. *Chem. Biol.* **2000**, *7*, 423.
(314) Blanchard, H.; Kodandapani, L.; Mittl, P. R. E.; Di Marco, S.; Krebs, J. F.; Wu, J. C.; Tomaselli, K. J.; Grütter, M. G. *Structure* **1999**, *7*, 1125.
(315) Riedl, S. J.; Fuentes-Prior, P.; Renatus, M.; Kairies, N.; Krapp, S.; Huber, R.; Salvesen, G. S.; Bode, W. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 14790.
(316) Rylander, P. N.; Tarbell, D. S. *J. Am. Chem. Soc.* **1950**, *72*, 3021.
(317) Xu, G.; Cirilli, M.; Huang, Y.; Rich, R. L.; Myszka, D. G.; Wu, H. *Nature (London)* **2001**, *410*, 494.
(318) Isenman, D. E. *Ann. N. Y. Acad. Sci.* **1983**, *421*, 277.
(319) Van Leuven, F.; Marynen, P.; Cassiman, J.-J.; Van den Berghe, H. *Biochem. J.* **1982**, *203*, 405.
(320) Sottrup-Jensen, L.; Petersen, T. E.; Magnusson, S. *FEBS Lett.* **1981**, *128*, 123.
(321) Dodds, A. W.; Ren, X. D.; Willis, A. C.; Law, S. K. A. *Nature (London)* **1996**, *379*, 177.
(322) Law, S. K. A.; Dodds, A. W. *Protein Sci.* **1997**, *6*, 263.
(323) Wang, D.; Yuan, A. I.; Feinman, R. D. *Biochemistry* **1984**, *23*, 2807.
(324) Nair, S. A.; Cooperman, B. S. *J. Biol. Chem.* **1998**, *273*, 17459.
(325) Shieh, B.-H.; Potempa, J.; Travis, J. *J. Biol. Chem.* **1989**, *264*, 13420.
(326) Schwartz, B. S.; España, F. *J. Biol. Chem.* **1999**, *274*, 15278.
(327) Olson, S. T.; Stephens, A. W.; Hirs, C. H. W.; Bock, P. E.; Björk, I. *J. Biol. Chem.* **1995**, *270*, 9717.
(328) Kvassman, J. O.; Verhamme, I.; Shore, J. D. *Biochemistry* **1998**, *37*, 7, 15491.
(329) Ohlsson, K.; Laurell, C. B. *Clin. Sci. Mol. Med.* **1976**, *51*, 87.
(330) Collen, D.; Wiman, B. In *The Physiological Inhibitors of Coagulation and Fibrinolysis*; Collen, D., Wiman, B., Verstraete, M., Eds.; Elsevier/North-Holland Biomedical Press: Amsterdam, 1979.
(331) Fuchs, H. E.; Shifman, M. A.; Pizzo, S. V. *Biochim. Biophys. Acta* **1982**, *716*, 151.
(332) Pizzo, S. V.; Mast, A. E.; Feldman, S. R.; Salvesen, G. *Biochim. Biophys. Acta* **1988**, *967*, 158.
(333) Pratt, C. W.; Church, F. C.; Pizzo, S. V. *Arch. Biochem. Biophys.* **1988**, *262*, 111.
(334) Gonias, S. L.; Fuchs, H. E.; Pizzo, S. V. *Thromb. Haemostasis* **1982**, *48*, 208.
(335) Perlmutter, D. H.; Glover, G. I.; Rivetna, M.; Schasteen, C. S.; Fallon, R. J. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 3753.
(336) Joslin, G.; Fallon, R. J.; Bullock, J.; Adams, S. P.; Perlmutter, D. H. *J. Biol. Chem.* **1991**, *266*, 11282.
(337) Schulze, A. J.; Huber, R.; Bode, W.; Engh, R. A. *FEBS Lett.* **1994**, *344*, 117.
(338) Maekawa, H.; Tollefsen, D. M. *J. Biol. Chem.* **1996**, *271*, 18604.
(339) Strickland, D. K.; Ashcom, J. D.; Williams, S.; Burgess, W. H.; Migliorini, M.; Argraves, W. S. *J. Biol. Chem.* **1990**, *265*, 17401.
(340) Kristensen, T.; Moestrup, S. K.; Gliemann, J.; Bendtsen, L.; Sand, O.; Sottrup-Jensen, L. *FEBS Lett.* **1990**, *276*, 151.
(341) Nykjaer, A.; Petersen, C. M.; Moller, B. K.; Jensen, P. H.; Moestrup, S. K.; Holtet, T. L.; Etzerodt, M.; Thøgersen, H. C.; Munch, M.; Andreasen, A. M. *J. Biol. Chem.* **1992**, *267*, 14543.
(342) Orth, K.; Madison, E. L.; Gething, M.-J.; Sambrook, J. F.; Herz, J. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 7422.
(343) Poller, W.; Willnow, T. E.; Hilpert, J.; Herz, J. *J. Biol. Chem.* **1995**, *270*, 2841.
(344) Kounnas, M. Z.; Church, F. C.; Argraves, W. S.; Strickland, D. K. *J. Biol. Chem.* **1996**, *271*, 6523.
(345) Kasza, A.; Petersen, H. H.; Heegaard, C. W.; Oka, K.; Christensen, A.; Dubin, A.; Chan, L.; Andreasen, P. A. *Eur. J. Biochem.* **1997**, *248*, 270.
(346) Storm, D.; Herz, J.; Trinder, P.; Loos, M. *J. Biol. Chem.* **1997**, *272*, 31043.
(347) Knauer, D. J.; Majumdar, D.; Fong, P.-C.; Knauer, M. F. *J. Biol. Chem.* **2000**, *275*, 37340.
(348) Herz, J.; Kowal, R. C.; Goldstein, J. L.; Brown, M. S. *EMBO J.* **1990**, *9*, 1769.
(349) Moestrup, S. K.; Holtet, T. L.; Etzerodt, M.; Thøgersen, H. C.; Nykjær, A.; Andreasen, P. A.; Rasmussen, H. H.; Sottrup-Jensen, L.; Gliemann, J. *J. Biol. Chem.* **1993**, *268*, 13691.
(350) Warshawsky, I.; Bu, G.; Schwartz, A. L. *Biochemistry* **1995**, *34*, 3404.
(351) Jeon, H.; Meng, W.; Takagi, J.; Eck, M. J.; Springer, T. A.; Blacklow, S. C. *Nat. Struct. Biol.* **2001**, *8*, 499.
(352) Springer, T. A. *J. Mol. Biol.* **1998**, *283*, 837.
(353) Chen, W.-J.; Goldstein, J. L.; Brown, M. S. *J. Biol. Chem.* **1990**, *265*, 3116.
(354) Fass, D.; Blacklow, S.; Kim, P. S.; Berger, J. M. *Nature (London)* **1997**, *388*, 691.
(355) Simonovic, M.; Dolmer, K.; Huang, W.; Strickland, D. K.; Volz, K.; Gettins, P. G. W. *Biochemistry* **2001**, *40*, 15127.
(356) Daly, N. L.; Djordjevic, J. T.; Kroon, P. A.; Smith, R. *Biochemistry* **1995**, *34*, 14474.
(357) Daly, N. L.; Scanlon, M. J.; Djordjevic, J. T.; Kroon, P. A.; Smith, R. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 6334.
(358) Huang, W.; Dolmer, K.; Gettins, P. G. W. *J. Biol. Chem.* **1999**, *274*, 14130.
(359) North, C. L.; Blacklow, S. C. *Biochemistry* **1999**, *38*, 3926.
(360) Dolmer, K.; Huang, W.; Gettins, P. G. W. *J. Biol. Chem.* **2000**, *275*, 3264.
(361) Beglova, N.; North, C. L.; Blacklow, S. C. *Biochemistry* **2001**, *40*, 2808.
(362) Strickland, D. K.; Kounnas, M. Z. *Trends Cardiovasc. Med.* **1997**, *7*, 9.
(363) Stefansson, S.; Lawrence, D. A.; Argraves, W. S. *J. Biol. Chem.* **1996**, *271*, 8215.
(364) Moestrup, S. K.; Gliemann, J.; Pallesen, G. *Cell Tissue Res.* **1992**, *269*, 375.
(365) Christensen, E. I.; Nielsen, S.; Moestrup, S. K.; Borre, C.; Maunsbach, A. B.; de Heer, E.; Ronco, P.; Hammond, T. G.; Verroust, P. *Eur. J. Cell Biol.* **1995**, *66*, 349.
(366) Wyne, K. L.; Pathak, K.; Seabra, M. C.; Hobbs, H. H. *Arterioscler., Thromb. Vasc. Biol.* **1996**, *16*, 407.
(367) Mast, A. E.; Enghild, J. J.; Pizzo, S. V.; Salvesen, G. *Biochemistry* **1991**, *30*, 1723.
(368) Stefansson, S.; Muhammad, S.; Cheng, X. F.; Battey, F. D.; Strickland, D. K.; Lawrence, D. A. *J. Biol. Chem.* **1998**, *273*, 6358.
(369) Malek, R.; Aulak, K. S.; Davis, A. E., III *Clin. Exp. Immunol.* **1996**, *105*, 191.
(370) de Smet, B. J.; de Boer, J. P.; Agterberg, J.; Rigter, G.; Bleeker, W. K.; Hack, C. E. *Blood* **1993**, *81*, 56.
(371) Nykjær, A.; Kjoller, L.; Cohen, R. L.; Lawrence, D. A.; Gliemann, J.; Andreasen, P. A. *Ann. NY Acad. Sci.* **1994**, *737*, 483.
(372) Moestrup, S. K.; Kaltoft, K.; Sottrup-Jensen, L.; Gliemann, J. *J. Biol. Chem.* **1990**, *265*, 12623.
(373) Herz, J.; Hamann, U.; Rogne, S.; Myklebost, O.; Gausepohl, H.; Stanley, K. K. *EMBO J.* **1988**, *7*, 4119.
(374) Dolmer, K.; Huang, W.; Gettins, P. G. W. *Biochemistry* **1998**, *37*, 17016.
(375) Lalazar, A.; Weisgraber, K. H.; Rall, S. C. J.; Giladi, H.; Innerarity, T. L.; Levanon, A. Z.; Boyles, J. K.; Amit, B.; Gorecki, M.; Mahley, R. W. *J. Biol. Chem.* **1988**, *263*, 3542.

(376) Wilson, C.; Wardell, M. R.; Weisgraber, K. H.; Mahley, R. W.; Agard, D. A. *Science* **1991**, *252*, 1817.
(377) Nielsen, K. L.; Holtet, T. L.; Etzerodt, M.; Moestrup, S. K.; Gliemann, J.; Sottrup-Jensen, L.; Thøgersen, H. C. *J. Biol. Chem.* **1996**, *271*, 12909.
(378) Rodenburg, K. W.; Kjoller, L.; Petersen, H. H.; Andreasen, P. A. *Biochem. J.* **1998**, *329*, 55.
(379) Knauer, M. F.; Hawley, S. B.; Knauer, D. J. *J. Biol. Chem.* **1997**, *272*, 12261.
(380) Knauer, M. F.; Crisp, R. J.; Kridel, S. J.; Knauer, D. J. *J. Biol. Chem.* **1999**, *274*, 275.
(381) Russell, D. W.; Brown, M. S.; Goldstein, J. L. *J. Biol. Chem.* **1989**, *264*, 21682.
(382) Rettenberger, P. M.; Oka, K.; Ellgaard, L.; Petersen, H. H.; Christensen, A.; Martensen, P. M.; Monard, D.; Etzerodt, M.; Chan, L.; Andreasen, P. A. *J. Biol. Chem.* **1999**, *274*, 8973.
(383) Andersen, O. M.; Petersen, H. H.; Jacobsen, C.; Moestrup, S. K.; Etzerodt, M.; Andreasen, P. A.; Thøgersen, H. C. *Biochem. J.* **2001**, *357*, 289.
(384) Nykjær, A.; Kjoller, L.; Cohen, R. L.; Lawrence, D. A.; Garni-Wagner, B. A.; Todd, R. F., III.; van Zonneveld, A.-J.; Gliemann, J.; Andreasen, P. A. *J. Biol. Chem.* **1994**, *269*, 25668.
(385) Brown, M. S.; Goldstein, J. L. *Science* **1986**, *232*, 34.
(386) Conese, M.; Olson, D.; Blasi, F. *J. Biol. Chem.* **1994**, *269*, 17886.
(387) Conese, M.; Nykjaer, A.; Petersen, C. M.; Cremona, O.; Pardi, R.; Andreasen, P. A.; Gliemann, J.; Christensen, E. I.; Blasi, F. *J. Cell Biol.* **1995**, *131*, 1609.
(388) Webb, D. J.; Nguyen, D. H. D.; Sankovic, M.; Gonias, S. L. *J. Biol. Chem.* **1999**, *274*, 7412.
(389) Nykjaer, A.; Conese, M.; Christensen, E. I.; Olson, D.; Cremona, O.; Gliemann, J.; Blasi, F. *EMBO J.* **1997**, *16*, 2610.
(390) Zhang, J. C.; Sakthivel, R.; Kniss, D.; Graham, C. H.; Strickland, D. K.; McCrae, K. R. *J. Biol. Chem.* **1998**, *273*, 32273.
(391) Knauer, M. F.; Kridel, S. J.; Hawley, S. B.; Knauer, D. J. *J. Biol. Chem.* **1997**, *272*, 29039.
(392) Li, Y.; Marzolo, M. P.; van Kerkhof, P.; Strous, G. J.; Bu, G. *J. Biol. Chem.* **2000**, *275*, 17187.
(393) Goretzki, L.; Mueller, B. M. *J. Cell Sci.* **1997**, *110*, 1395.
(394) Sakthivel, R.; Zhang, J.-C.; Strickland, D. K.; Gåfvels, M.; McCrae, K. R. *J. Biol. Chem.* **2001**, *276*, 555.
(395) Goretzki, L.; Mueller, B. M. *Biochem. J.* **1998**, *336*, 381.
(396) Misra, U. K.; Chu, C. T.-C.; Gawdi, G.; Pizzo, S. V. *J. Biol. Chem.* **1994**, *269*, 12541.
(397) Kinoshita, A.; Whelan, C. M.; Smith, C. J.; Mikhailenko, I.; Rebeck, G. W.; Strickland, D. K.; Hyman, B. T. *J. Neurosci.* **2001**, *21*, 8354.
(398) Gotthardt, M.; Trommsdorff, M.; Nevitt, M. F.; Shelton, J.; Richardson, J. A.; Stockinger, W.; Nimpf, J.; Herz, J. *J. Biol. Chem.* **2000**, *275*, 25616.
(399) Brown, M. S.; Herz, J.; Goldstein, J. L. *Nature (London)* **1997**, *388*, 629.
(400) Davis, C. G.; Goldstein, J. C.; Sudhof, T. C.; Anderson, R. G. W.; Russell, D. W.; Brown, M. S. *Nature (London)* **1987**, *326*, 760.
(401) Mikhailenko, I.; Considine, W.; Argraves, K. M.; Loukinov, D.; Hyman, B. T.; Strickland, D. K. *J. Cell Sci.* **1999**, *112*, 3269.
(402) Crisp, R. J.; Knauer, D. J.; Knauer, M. F. *J. Biol. Chem.* **2000**, *275*, 19628.
(403) Kurdowska, A.; Travis, J. *J. Biol. Chem.* **1990**, *265*, 21023.
(404) Schuster, M. G.; Enriquez, P. M.; Curran, P.; Cooperman, B. S.; Rubin, H. *J. Biol. Chem.* **1992**, *267*, 5056.
(405) Banda, M. J.; Rice, A. G.; Griffin, G. L.; Senior, R. M. *J. Exp. Med.* **1988**, *167*, 1608.
(406) Dawson, D. W.; Volpert, O. V.; Gillis, P.; Crawford, S. E.; Xu, H.-J.; Benedict, W.; Bouck, N. P. *Science* **1999**, *285*, 245.
(407) Stellmach, V.; Crawford, S. E.; Zhou, W.; Bouck, N. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 2593.
(408) Steele, F. R.; Chader, G. J.; Johnson, L. V.; Tombran-Tink, J. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 1526.
(409) Aymerich, M. S.; Alberdi, E. M.; Martinez, A.-M.; Becerra, S. P. *Invest. Ophthalmol. Vis. Sci.* **2001**, *42*, 3287.
(410) Corbin, L. W.; Church, F. C.; Hoffman, M. *Thromb. Res.* **1990**, *57*, 77.
(411) Church, F. C.; Pratt, C. W.; Hoffman, M. *J. Biol. Chem.* **1991**, *266*, 704.
(412) Schechter, I.; Berger, A. *Biochem. Biophys. Res. Commun.* **1967**, *27*, 157.
(413) Bao, J. J.; Reed-Fourquet, L.; Sifers, R. N.; Kidd, V. J.; Woo, S. *Genomics* **1988**, *2*, 165.
(414) Schechter, N. M.; Jordan, L. M.; James, A. M.; Cooperman, B. S.; Wang, Z. M.; Rubin, H. *J. Biol. Chem.* **1993**, *268*, 23626.
(415) Janciauskiene, S.; Rubin, H.; Lukacs, C. M.; Wright, H. T. *J. Biol. Chem.* **1998**, *273*, 28360.
(416) Hsieh, M.-C.; Cooperman, B. S. *Biochemistry* **2002**, *41*, 2990.
(417) Pratt, C. W.; Whinna, H. C.; Church, F. C. *J. Biol. Chem.* **1992**, *267*, 8789.
(418) Stump, D. C.; Thienpont, M.; Collen, D. *J. Biol. Chem.* **1986**, *261*, 12759.
(419) Meijers, J. C. M.; Kanters, D. H. A. J.; Vlooswijk, R. A. A.; vanErp, H. E.; Hessing, M.; Bouma, B. N. *Biochemistry* **1988**, *27*, 4231.
(420) Hermans, J. M.; Jones, R.; Stone, S. R. *Biochemistry* **1994**, *33*, 5440.
(421) Rezaie, A. R.; Cooper, S. T.; Church, F. C.; Esmon, C. T. *J. Biol. Chem.* **1995**, *270*, 25336.
(422) Stein, P. E.; Tewkesbury, D. A.; Carrell, R. W. *Biochem. J.* **1989**, *262*, 103.
(423) Poulsen, K.; Haber, E.; Burton, J. *Biochim. Biophys. Acta* **1976**, *452*, 533.
(424) Frazer, J. K.; Jackson, D. G.; Gaillard, J. P.; Lutter, M.; Liu, Y. J.; Banchereau, J.; Capra, J. D.; Pascual, V. *Eur. J. Immunol.* **2000**, *30*, 3039.
(425) Remold-O'Donnell, E.; Chin, J.; Alberts, M. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 5635.
(426) Cooley, J.; Takayama, T. K.; Shapiro, S. D.; Schechter, N. M.; Remold-O'Donnell, E. *Biochemistry* **2001**, *40*, 15762.
(427) Ye, R. D.; Wun, T.; Sadler, J. E. *J. Biol. Chem.* **1987**, *262*, 3718.
(428) Antalis, T. M.; Clark, M. A.; Barnes, T.; Lehrbach, P. R.; Devine, P. L.; Schevzov, G.; Goss, N. H.; Stephens, R. W.; Tolstoshev, P. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 985.
(429) Schick, C.; Kamachi, Y.; Bartuski, A. J.; Çataltepe, S.; Schechter, N. M.; Pemberton, P. A.; Silverman, G. A. *J. Biol. Chem.* **1997**, *272*, 1849.
(430) Sheng, S.; Truong, B.; Fredrickson, D.; Wu, R.; Pardee, A. B.; Sager, R. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 499.
(431) Coughlin, P.; Sun, J. R.; Cerruti, L.; Salem, H. H.; Bird, P. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 9417.
(432) Morgenstern, K. A.; Sprecher, C.; Holth, L.; Foster, D.; Grant, F. J.; Ching, A.; Kisiel, W. *Biochemistry* **1994**, *33*, 3432.
(433) Scott, F. L.; Hirst, C. E.; Sun, J.; Bird, C. H.; Bottomley, S. P.; Bird, P. I. *Blood* **1999**, *93*, 2089.
(434) Sun, J. R.; Coughlin, P.; Salem, H. H.; Bird, P. *Biochim. Biophys. Acta* **1995**, *1252*, 28.
(435) Inagi, R.; Miyata, T.; Suzuki, D.; Toyoda, M.; Wada, T.; Ueda, Y.; Izuhara, Y.; Sakai, H.; Nangaku, M.; Kurokawa, K. *Biochem. Biophys. Res. Commun.* **2001**, *286*, 1098.
(436) Dahlen, J. R.; Foster, D. C.; Kisiel, W. *Biochemistry* **1997**, *36*, 14874.
(437) Sun, J.; Bird, C. H.; Sutton, V.; McDonald, L.; Coughlin, P. B.; De Jong, T. A.; Trapani, J. A.; Bird, P. I. *J. Biol. Chem.* **1996**, *271*, 27802.
(438) Riewald, M.; Schleef, R. R. *J. Biol. Chem.* **1995**, *270*, 26754.
(439) Askew, Y. S.; Pak, S. C.; Luke, C. J.; Askew, D. J.; Cataltepe, S.; Mills, D. R.; Kato, H.; Lehoczky, y. J.; Dewar, K.; Birren, B.; Silverman, G. A. *J. Biol. Chem.* **2001**, *276*, 49320.
(440) Abts, H. F.; Welss, T.; Mirmohammadsadegh, A.; Kohrer, K.; Michel, G.; Ruzicka, T. *J. Mol. Biol.* **1999**, *293*, 29.
(441) Sheehan, J. P.; Tollefsen, D. M.; Sadler, J. E. *J. Biol. Chem.* **1994**, *269*, 32747.
(442) Lawrence, D. A.; Strandberg, L.; Ericson, J.; Ny, T. *J. Biol. Chem.* **1990**, *265*, 20293.
(443) Keijer, J.; Linders, M.; Wegman, J. J.; Ehrlich, H. J.; Mertens, K.; Pannekoek, H. *Blood* **1991**, *78*, 1254.
(444) Stefansson, S.; Petitclerc, E.; Wong, M. K. K.; McMahon, G. A.; Brooks, P. C.; Lawrence, D. A. *J. Biol. Chem.* **2001**, *276*, 8135.
(445) Scott, R. W.; Bergman, B. L.; Bajpai, A.; Hersh, R. T.; Rodriguez, H.; Jones, B. N.; Barreda, C.; Watts, S.; Baker, J. B. *J. Biol. Chem.* **1985**, *260*, 7029.
(446) Rovelli, G.; Stone, S. R.; Guidolin, A.; Sommer, J.; Monard, D. *Biochemistry* **1992**, *31*, 3542.
(447) Becerra, S. P.; Sagasti, A.; Spinella, P.; Notario, V. *J. Biol. Chem.* **1995**, *270*, 25992.
(448) Stratikos, E.; Alberdi, E.; Gettins, P. G. W.; Becerra, S. P. *Protein Sci.* **1996**, *5*, 2575.
(449) Sim, R. B.; Arlaud, G.; Colomb, M. *Biochim. Biophys. Acta* **1980**, *612*, 433.
(450) Schapira, M.; deAgostini, A.; Colman, R. W. *J. Clin. Invest.* **1982**, *69*, 462.
(451) van der Graaf, F.; Koedam, J. A.; Griffin, J. H.; Bouma, B. N. *Biochemistry* **1983**, *22*, 4860.
(452) Clarke, E. P.; Sanwal, B. D. *Biochim. Biophys. Acta* **1992**, *1129*, 246.
(453) Nagai, N.; Tetuya, Y.; Hosokawa, N.; Nagata, K. *Gene* **1999**, *227*, 241.
(454) Ikegawa, S.; Sudo, K.; Okui, K.; Nakamura, Y. *Cytogenet. Cell Genet.* **1995**, *71*, 182.
(455) Hattori, T.; Takahash, K.; Yutani, Y.; Fujisawa, T.; Nakanishi, T.; Takigawa, M. *J. Bone Miner. Metab.* **2000**, *18*, 328.
(456) Hastings, G. A.; Coleman, T. A.; Haudenschild, C. C.; Stefansson, S.; Smith, E. P.; Barthlow, R.; Cherry, S.; Sandkvist, M.; Lawrence, D. A. *J. Biol. Chem.* **1997**, *272*, 32474.
(457) Osterwalder, T.; Cinelli, P.; Baici, A.; Pennella, A.; Krueger, S. R.; Schrimpf, S. P.; Meins, M.; Sonderegger, P. *J. Biol. Chem.* **1998**, *273*, 2312.
(458) Xiao, G.; Liu, Y. E.; Gentz, R.; Sang, Q. A.; Ni, J.; Goldberg, I. D.; Shi, Y. E. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 3700.

(459) Ozaki, K.; Nagata, M.; Suzuki, M.; Fujiwara, T.; Miyoshi, Y.; Ishikawa, O.; Ohigashi, H.; Imaoka, S.; Takahashi, E.; Nakamura, Y. *Genes Chromosomes Cancer* **1998**, *22*, 179.
(460) Chang, W. S.; Chang, N. T.; Lin, S. C.; Wu, C. W.; Wu, F. Y. *Genes Chromosomes Cancer* **2000**, *29*, 240.
(461) Löbermann, H.; Tokuoka, R.; Deisenhofer, J.; Huber, R. *Protein Eng.* **1989**, *2*, 407.
(462) Song, H. K.; Lee, K. N.; Kwon, K.-S.; Yu, M.-H.; Suh, S. W. *FEBS Lett.* **1995**, *377*, 150.
(463) Ryu, S.-E.; Choi, H.-J.; Kwon, K.-S.; Lee, K. N.; Yu, M.-H. *Structure* **1996**, *4*, 1181.
(464) Baumann, U.; Huber, R.; Bode, W.; Grosse, D.; Lesjak, M.; Laurell, C. B. *J. Mol. Biol.* **1991**, *218*, 595.
(465) Lukacs, C. M.; Rubin, H.; Christianson, D. W. *Biochemistry* **1998**, *37*, 3297.
(466) Baumann, U.; Bode, W.; Huber, R.; Travis, J.; Potempa, J. *J. Mol. Biol.* **1992**, *226*, 1207.
(467) Stein, P. E.; Leslie, A. G. W.; Finch, J. T.; Carrell, R. W. *J. Mol. Biol.* **1991**, *221*, 941.
(468) Mourey, L.; Samama, J. P.; Delarue, M.; Petitou, M.; Choay, J.; Moras, D. *J. Mol. Biol.* **1993**, *232*, 223.
(469) Aertgeerts, K.; De Bondt, H. L.; De Ranter, C. J.; Declerck, P. J. *Nat. Struct. Biol.* **1995**, *2*, 891.
(470) Stout, T. J.; Graham, H.; Buckley, D. I.; Matthews, D. J. *Biochemistry* **2000**, *39*, 8460.
(471) Sharp, A. M.; Stein, P. E.; Pannu, N. S.; Carrell, R. W.; Berkenpas, M. B.; Ginsburg, D.; Lawrence, D. A.; Read, R. J. *Structure* **1999**, *7*, 110.
(472) Nar, H.; Bauer, M.; Stassen, J.; Lang, D.; Gils, A.; Declerck, P. J. *J. Mol. Biol.* **2000**, *297*, 683.
(473) Briand, C.; Kozlov, G.; Sonderegger, P.; Grütter, J. G. *FEBS Lett.* **2001**, *505*, 18.
(474) Chaillan-Huntington, C. E.; Patston, P. A. *J. Biol. Chem.* **1998**, *273*, 4569.
(475) Schechter, N. M.; Sprows, J. L.; Schoenberger, O. L.; Lazarus, G. S.; Cooperman, B. S.; Rubin, H. *J. Biol. Chem.* **1989**, *264*, 21308.
(476) Lawrence, D. A.; Olson, S. T.; Palaniappan, S.; Ginsburg, D. *J. Biol. Chem.* **1994**, *269*, 27657.
(477) Perry, D. J.; Harper, P. L.; Fairham, S.; Daly, M.; Carrell, R. W. *FEBS Lett.* **1989**, *254*, 174.
(478) Davis, A. E., III.; Aulak, K.; Parad, R. B.; Stecklein, H. P.; Eldering, E.; Hack, C. E.; Kramer, J.; Strunk, R. C.; Bissler, J.; Rosen, F. S. *Nat. Genet.* **1992**, *1*, 354.
(479) Skriver, K.; Wikoff, W. R.; Patston, P. A.; Tausk, F.; Schapira, M.; Kaplan, A. P.; Bock, S. C. *J. Biol. Chem.* **1991**, *266*, 9216.
(480) Siddique, Z. M.; McPhaden, A. R.; Whaley, K. *Clin. Exp. Immunol.* **1991**, *86*, 11.
(481) Levy, N. J.; Ramesh, N.; Cicardi, M.; Harrison, R. A.; Davis, A. E. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 265.
(482) Aulak, K. S.; Eldering, E.; Hack, C. E.; Lubbers, Y. P. T.; Harrison, R. A.; Mast, A.; Cicardi, M.; Davis, A. E., III. *J. Biol. Chem.* **1993**, *268*, 18088.
(483) Hopkins, P. C. R.; Crowther, D. C.; Carrell, R. W.; Stone, S. R. *J. Biol. Chem.* **1995**, *270*, 11866.
(484) Rubin, H.; Wang, Z. M.; Nickbarg, E. B.; McLarney, S.; Naidoo, N.; Schoenberger, O. L.; Johnson, J. L.; Cooperman, B. S. *J. Biol. Chem.* **1990**, *265*, 1199.
(485) Rubin, H.; Plotnick, M.; Wang, Z.; Liu, X.; Zhong, Q.; Schechter, N. M.; Cooperman, B. S. *Biochemistry* **1994**, *33*, 7627.
(486) Esmon, C. T. *Science* **1987**, *235*, 1348.
(487) Phillips, J. E.; Cooper, S. T.; Potter, E. E.; Church, F. C. *J. Biol. Chem.* **1994**, *269*, 16696.
(488) Cooper, S. T.; Church, F. C. *Biochim. Biophys. Acta* **1995**, *1246*, 29.
(489) Rosenberg, R. D. *Fed. Proc.* **1977**, *36*, 10.
(490) Chuang, Y.-J.; Gettins, P. G. W.; Olson, S. T. *J. Biol. Chem.* **1999**, *274*, 28142.
(491) Church, F. C.; Noyes, C. M.; Griffith, M. J. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 6431.
(492) Derechin, V. M.; Blinder, M. A.; Tollefsen, D. M. *J. Biol. Chem.* **1990**, *265*, 5623.
(493) Ehrlich, H. J.; Gebbink, R. K.; Keijer, J.; Linders, M.; Preissner, K. T.; Pannekoek, H. *J. Biol. Chem.* **1990**, *265*, 13029.
(494) Sherman, P. M.; Lawrence, D. A.; Yang, A. Y.; Vandenberg, E. T.; Paielli, D.; Olson, S. T.; Shore, J. D.; Ginsburg, D. *J. Biol. Chem.* **1992**, *267*, 7588.
(495) Holmes, W. E.; Lijnen, H. R.; Collen, D. *Biochemistry* **1987**, *26*, 5133.
(496) Eldering, E.; Huijbregts, C. C. M.; Lubbers, Y. T. P.; Longstaff, C.; Hack, C. E. *J. Biol. Chem.* **1992**, *267*, 7013.
(497) Skriver, K.; Radziejewska, E.; Silbermann, J. A.; Donaldson, V. H.; Bock, S. C. *J. Biol. Chem.* **1989**, *264*, 3066.
(498) Hermans, J. M.; Stone, S. R. *Biochem. J.* **1993**, *295*, 239.
(499) Evans, D. L.; McGrogan, M.; Scott, R. A.; Carrell, R. W. *J. Biol. Chem.* **1991**, *266*, 22307.
(500) Rezaie, A. R. *Biochemistry* **1999**, *38*, 14592.
(501) McGrath, M. E.; Klaus, J. L.; Barnes, M. G.; Bromme, D. *Nat. Struct. Biol.* **1997**, *4*, 105.
(502) Riedl, S. J.; Renatus, M.; Schwarzenbacher, R.; Zhou, Q.; Sun, C.; Fesik, S. W.; Liddington, R. C.; Salvesen, G. S. *Cell* **2001**, *104*, 791.

CR010170+