Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
function, and requirement for protease activation. The E1 glycoprotein may represent a new class of glycoproteins resembling in some ways cellular glycoproteins associated with the plasma membrane, and in other ways, cellular glycoproteins associated with the Golgi apparatus.

Acknowledgements
We thank Drs J. Armstrong, M. Lai, H. Niemann and P. Rottier who shared their unpublished data with us. Studies in the authors' laboratories were supported by grants GM 31698 and AI 18997 from the NIH. The opinions expressed are the private views of the authors and should not be construed as official or necessarily reflecting the view of the Uniformed Services School of Medicine or the Department of Defense.

References
1 Siddell, S. G., Anderson, R., Cavanagh, D., Fajjwara, K., Klenk, H. D., Macnaughton, M. R., Pensaert, M., Stohlman, S. A., Sturman, L. and van der Zeijst, B. A. M. (1982) Nature 306, 751-752
2 Rottier, P. J. M., van der Zeijst, B. A. M., Spaan, W. J. M. and Horznéznék, M. C. (eds) (1984) Adv. Virus Res. 35 (Pergamon Press 44, 804-812)
3 Siddell, S., Wege, H. and ter Meulen, V. (1983) Curr. Top. Microbiol. Immunol., 102, 1-20
4 Rottier, P. J. M., van der Zeijst, B. A. M., Spaan, W. J. M. and Horznéznék, M. C. (eds) (1984) Adv. Exp. Med. Biol. 173, 1-409
5 ter Meulen, V., Siddell, S. and Wege, H. (eds) (1981) Adv. Exp. Med. Biol. 142, 1-38
6 Talbot, P. J., Salmi, A. A., Knobler, R. L. and Buchmeier, M. J. (1984) Virology 132, 250-260
7 Brayton, P. R., Lai, M. M. C., Patton, C. D. and Stohlman, S. A. (1982) J. Virol. 42, 845-853
8 Lai, M. M. C., Patton, C. D. and Stohlman, S. A. (1982) J. Virol. 44, 488-492
9 Lai, M. M. C., Brayton, P. R., Arnen, R. C., Patton, C. C., Pugh, C. and Stohlman, S. A. (1981) J. Virol. 39, 823-834
10 Lai, M. M. C., Patton, C. C., Baric, R. S. and Stohlman, S. A. (1983) J. Virol. 46, 1027-1033
11 Lai, M. M. C., Baric, R. S., Brayton, P. R. and Stohlman, S. A. (1984) Proc. Natl Acad. Sci. USA 81, 3626-3630
12 Spaan, W., Delius, H., Skinner, M., Armstrong, J., Rottier, P., Smeekens, S., van der Zeijst, B. A. M. and Siddell, S. G. (1983) EMBO J. 2, 1839-1844
13 Baric, R. S., Stohlman, S. A. and Lai, M. M. C. (1983) J. Virol. 48, 633-640
14 Siddell, S., Wege, H., Barthei, A. and ter Meulen, V. (1981) in Biochemistry and Biology of Coronavirus (eds ter Meulen, V., Siddell, S. and Wege, H.), pp. 193-208, Plenum Press
15 Rottier, P. J. M., Spaan, W. J. M., Horznéznék, M. and van der Zeijst, B. A. M. (1981) J. Virol. 30, 20-26
16 Leibowitz, J. L., Weitz, S. R., Paavola, E. and Bond, C. W. (1982) J. Virol. 45, 905-913
17 Rottier, P. J. M., Horznéznék, M. C. and van der Zeijst, B. A. M. (1981) J. Virol. 40, 350-357
18 Niemann, H. and Klenk, H. D. (1981) J. Mol. Biol. 153, 993-1010
19 Holmes, K. V., Doffer, E. W. and Behnke, J. N. (1981) in Biochemistry and Biology of Coronaviruses (eds Meulen, V., Siddell, S. and Wege, H.), pp. 133-142, Plenum Press
20 Sturman, L. S. and Holmes, K. V. (1984) in Molecular Biology and Pathogenesis of Coronavirus (editors P. J. M. van der Zeijst, B. A. M., Spaan, W. J. M. and Horznéznék, M. C.), pp. 25-35, Plenum Press
21 Sturman, L. S. (1981) in Biochemistry and Biology of Coronavirus (eds Meulen, V., Siddell, S. and Wege, H.), pp. 1-18, Plenum Press
22 Rottier, P., Brandenburg, D., Armstrong, J., and van der Zeijst, B. and Warren, G. (1984) Proc. Natl Acad. Sci. USA 81, 1421-1425
23 Armstrong, J., Niemann, H., Smeekens, S., Rottier, P. and Warren, G. (1984) Nature 308, 751-752
24 Niemann, H., Geyer, B., Klenk, H. D., Linder, D., Sturm, S. and Wirth, M. (1984) EMBO J. 3, 665-670
25 Toozé, J., Toozé, S. and Warren, G. (1984) Eur. J. Cell Biol. 33, 281-293
26 Stern, D. F. and Selton, B. M. (1982) J. Virol. 44, 804-812
27 Boursnell, M. E. G., Brown, T. K. D. and Binn, M. M. (1984) Virus Research 1, 303-313
28 Hegeness, M. H., Smith, P. R. and Choppin, P. W. (1982) Proc. Natl Acad. Sci. USA 79, 6232-6236

α1-Antitrypsin and the serpins: variation and countervariation

Robin Carroll and James Travis

α1-Antitrypsin is a plasma protein which protects elastic tissue from proteolytic attack. Consequently, genetic deficiency, or the oxidation of its reactive centre in cigarette smokers can result in the degenerative lung disease emphysema. Structural studies explain the mechanisms involved and have also drawn attention to a new family of serine protease inhibitors. The specificity of each of these inhibitors is primarily dependent on a single amino acid at its reactive centre. Site-directed mutagenesis is enabling the production of specifically designed inhibitors for therapeutic use, including an improved replacement for α1-antitrypsin deficiency.

Human plasma contains several inhibitors of proteolytic enzymes that together form some 10% of its protein content. A major example is the broad-spectrum inhibitor, α2-macroglobulin, previously described in TIBS (James, 1980). How-ever, the best studied of the inhibitors, and one present in greatest concentration in plasma, is the more specifically targeted α1-antitrypsin (also called α1-proteinase inhibitor). Although this will inhibit most of the serine proteinases, it is really an anti-elastase and its prime physiological task is the inhibition of elastase released by neutrophil leucocytes.

The function of α1-antitrypsin is known because its genetic deficiency leads to a premature breakdown of connective tissue, to give a loss of elasticity in the lungs – a condition known as emphysema. Extensive studies on the molecular pathology of this genetic deficiency culminated recently in the completion of the structure of the normal and variant molecules.

An unexpected bonus from these structural studies is that α1-antitrypsin proved to be the archetype of a new superfamily of homologous proteins. We have called this family, for convenience, the serpins since it is primarily a group of Serine Proteinase Inhibitors.

It is becoming clear that the lessons learnt from α1-antitrypsin are also applicable to the other members of the serpin family. One important conclusion, already apparent, is that the inhibitory specificity of the serpins is in each case primarily defined by a single amino acid at the reactive centre of the molecule. This is strongly supported by previous work on the unrelated proteinase inhibitors (Table 1). An exciting corollary is the ability to redesign the specificity of inhibition by a single substitution at the reactive centre. This ability to design inhibitors for a specific purpose is already being developed, for therapeutic purposes, using simple recombinant-DNA modifications of α1-antitrypsin.

Robin Carroll is at the Pathology Department, Christchurch Clinical School of Medicine, Christchurch Hospital, New Zealand and James Travis was at this address on leave from the Biochemistry Department, The University of Georgia, Athens, GA, USA. In August 1984, R. Carrell was awarded the Pharmacia Prize by the New Zealand Biological Society. His presentation lecture covered material presented in this review.

© 1985, Elsevier Science Publishers B.V., Amsterdam 072-5785/85/$02.00.
Table I. Comparison of the reactive centres of the divergent serpins with the unrelated but homologous plant inhibitors, showing that specificity depends on the P₁ residue

| Inhibitor     | Target     | Reactive centre residues |
|---------------|------------|--------------------------|
| α₁-Antitrypsin| elastase   | P₁ P₁⁺ P₁⁻ P₁⁻⁺ P₁⁻⁻     |
| Valine mutant | elastase   | Pro Met Ser Ile Pro Pro  |
| Pittsburgh mutant | thrombin    | Gly Ile Ser Leu Asn Pro  |
| α₁-Antichymotrypsin | chymase  | Leu Ser Ala Leu Val  |
| Mouse contrasin | trypsin'    | Arg Lys Ala Ile Leu Pro  |
| CI-Inhibitor  | kallikrein  | Ala Arg Thr Leu Leu Val  |
| Ovalbumin    | ? elastase | Ala Ser Val Ser Glu Pro  |
| Angiotensinogen | ? function | Pro Gly Ser Pro Glu Pro  |

**Plant inhibitors**

| Soybean      | trypsin    | Thr Lys Ser Asn Pro Pro  |
| Garden bean  | elastase   | Thr Ala Ser Ile Pro Pro  |
| Lima bean    | trypsin    | Thr Lys Ser Ile Pro Pro  |

**Structure and function**

α₁-Antitrypsin is a relatively small glycoprotein of Mr 51 000, with 394 amino acid residues and three oligosaccharide side chains. It readily diffuses into the tissue spaces where it forms 1:1 complexes with its target proteases, principally neutrophil elastase. The complex of inhibitor and enzyme is subsequently removed from the circulation and catabolized in the liver and spleen.

The inhibitor is thought to trap the protease by presenting its reactive site as an ideal substrate. There is rapid association between the inhibitor and enzyme, but negligible dissociation, with consequent formation of a tightly bound enzyme-inhibitor complex.

The reactive site of α₁-antitrypsin is situated near its C-terminus, centred on a methionine residue at position 358. Complex formation with elastase is accompanied by proteolytic attack on the bond between this methionine and the adjacent serine at position 359. The key role of methionine 358 in defining the specificity of inhibition is indicated in Table I. For the serpins in general, the reactive centre, or P₁ residue acts as a bait for the appropriate serine proteinase with an affinity that is familiar to protein chemists, i.e. methionine (or valine) for elastase, lysine for trypsin, arginine for thrombin.

This descriptive concept of the inhibitor as a trap, with the active centre acting as a bait, is supported by the crystallographic structure of α₁-antitrypsin recently determined by Huber and colleagues. They were unable to crystallize the native α₁-antitrypsin but readily crystallized the cleaved inhibitor released after complex formation. As expected, the structure of α₁-antitrypsin is highly ordered, with some 40% β-sheet and 30% α-helix. This tight structure is presumably designed to firmly fix the reactive centre in an exposed site, and in its optimal 'substrate' configuration. This is a presumption because in the cleaved structure the reactive centre inhibitor and avoids a misleading conclusion to a target enzyme. One embarrassing example of misnaming is α₁-antitrypsin itself, which is not primarily an inhibitor of trypsin. Similarly, α₁-antichymotrypsin is not primarily an inhibitor of chymotrypsin, but rather of cathepsin G and the chymases. The ambiguity of the historical nomenclature is clearly illustrated by the recent recognition of a homologue, in the mouse, of human α₁-antichymotrypsin. This mouse α₁-antichymotrypsin has lysine rather than leucine at its reactive site, making it a Lyserpin rather than a Leuserpin. Hence, it does not inhibit chymotrypsin but is a highly effective inhibitor of trypsin-like proteases.

The ability to recognize the reactive centre by homology has also allowed the identification of putative inhibitory activity in two unexpected members of the serpin family. These are ovalbumin and angiotensinogen, neither of which are known to have an inhibitory function. However, both retain homology about the reactive site – perhaps ovalbumin is an Alaserpin and angiotensinogen is a Glyserpin?

Angiotensinogen is a particularly interesting addition to the family. It has a recognized physiological role as the source of the small peptides that control blood pressure: angiotensin I and II.
These are cleaved as a ten-residue segment from the N-terminus of the molecule. This activity of angiotensinogen has highlighted the likelihood that in the serpins in general, the N-terminal tail of the molecule forms a second functional site. This tail is of variable length in the individual members of the family; it does not have ordered structure and is on the exterior of the molecule where it is accessible to cleavage. It is this N-terminal tail that is the likely candidate for the major heparin binding site in antithrombin. In α₁-antitrypsin, N-terminal cleavage is a significant cause of the microheterogeneity of the molecule. It is not known whether this cleavage affects the function of the inhibitor or whether the released peptides function as physiological messengers.

Mutation, deficiency and disease

Two genetic variations of the amino acid sequence of α₁-antitrypsin are of particular importance; both are confined to Europeans and result in a decreased plasma concentration of the inhibitor. These are the S and Z deficiency variants which differ from the normal M α₁-antitrypsin in the substitution of two different glutamic acids. The product of the S deficiency gene (264 Glu→Val) is present in about 60% of the concentration of the normal M gene, and that of the Z gene (342 Glu→Lys) at 15%. The two glutamic acids substituted in these variants both form important salt bridges which normally stabilize the molecule and it is not surprising that both variants have a decreased stability, as measured by heat stress. The loss of stability of the S variant is compatible with an increased turnover of the nascent S polypeptide but the prime cause of its decreased production is likely to be the presence of an aberrant intron splice site introduced by the point mutation.

The more severe Z deficiency is of great interest as it involves a defect in the pathway of processing and secretion of the polypeptide. Electron microscopy shows a gross accumulation of Z α₁-antitrypsin at the level of the rough-to-smooth endoplasmic junction (Fig. 3). Thus, there is a partial blockage of incompletely processed α₁-antitrypsin just before it enters the Golgi vesicles. The cause of this blockage still has to be demonstrated but experience with a similar defect in immunoglobulin processing suggests blockage of transport, rather than a failure of oligosaccharide processing.

One in ten ZZ homozygotes will have liver disease, usually in early infancy. Whether this results directly from accumulation of the Z protein in the liver or is a consequence of the plasma deficiency of inhibitor is still being debated. What is clear, however, is that the lung disease is directly related to the plasma deficiency.

Smoking and emphysema

In the course of a lifetime, there is a gradual and irreversible loss of elasticity of the lungs. Up to the age of thirty, the lungs have an elastic volume of 3.5 litres but this decreases through later life by a mean value of 35 ml per year. Consequently, by the end of their lifespan, most people will have some limitation of lung function, and since there is considerable individual variation there will be some who will have lost a sufficient proportion of their elastic volume to give the crippling lung disease emphysema. The chances of getting emphysema are greatly increased if there is an accelerated loss of capacity. A number of factors are likely to be involved but two documented occurrences of accelerated loss are in heavy smokers and in people with severe α₁-antitrypsin deficiency; both groups have a mean annual loss of 80 ml per year. Again, there is considerable individual variation but the onset of emphysema, which is an unlikely event in the normal individual, becomes a likely event in the heavy smoker and also in the individual with homozygous α₁-antitrypsin deficiency. However, it becomes a certain event in the α₁-antitrypsin deficient individual who also smokes - here the loss of expiratory volume increases to

Fig. 2. (a) The reactive centre at α₁-antitrypsin is exposed on a strained loop probably hinged on residue 342 (site of the Z mutation). (b) Cleavage by a target enzyme releases the loop to give the stable post-complex form with Met 358 separated by 69Å from Ser 359.

The experimental results fit well with the histological changes. The RNA message (mRNA) for the M and Z proteins can be readily extracted from human liver and shown to be present in equivalent amounts. Addition of the mRNAs to cell-free systems shows equal translation and addition of oligosaccharide side chains for M and Z.

This is confirmed by injecting the mRNAs into the Xenopus oocyte; however, although there is equal production and processing, there is decreased secretion of the Z antitrypsin.

Whatever the mechanism involved, the end result is the same - a significant deficiency of plasma α₁-antitrypsin. About 4% of Europeans are heterozygotes for the Z defect, i.e. genotype MZ, and nearly one in a thousand will have the severe ZZ or SZ deficiency.

The homozygous ZZ individual is liable to develop liver and lung disease.

Fig. 3. Electron micrograph of liver from a ZZ patient. The Z antitrypsin accumulates as large amorphous granules in the endoplasmic reticulum.
300 ml per year and most in this group will have emphysema by the age of forty and be dead before they are fifty-five.\textsuperscript{1-5}

This acceleration of the onset of emphysema is readily explained in terms of the current belief\textsuperscript{24} that the elasticity of the lung is maintained by a balance between the release of destructive enzymes (principally elastase) and the protective effect of inhibitors (principally α1-antitrypsin). The balance between protease and antiprotease is destroyed by the genetic deficiency of α1-antitrypsin. Genetic deficiency in combination with habitual smoking is double jeopardy, since smoking results in an increased concentration of leucocytes in the lung and consequently increased exposure to neutrophil elastase.

In addition, smokers are subject to another risk factor, oxidative stress\textsuperscript{25-28}, that inactivates antitrypsin. The investigation of this oxidative inactivation has revealed a subtle control mechanism that is built into the reactive site of α1-antitrypsin. The realization that oxidation can inactivate α1-antitrypsin has also led to an understanding of a number of lung pathologies, in particular the emphysema of the heavy cigarette smoker.\textsuperscript{24}

The reactive centre methionine of α1-antitrypsin fits the active-site pocket of elastase and, therefore, functions as a bait for its entrapment. However, methionine is readily oxidizable to methionine sulphoxide, which has a considerably larger bulk and does not readily fit into the elastase pocket. As a consequence, oxidized α1-antitrypsin is an ineffective inhibitor of neutrophil elastase. This provides a real problem as there are several situations where oxidative stress occurs, particularly in the exposed areas of the lung, sufficient to give demonstrable inactivation of α1-antitrypsin. This occurs with direct exposure to oxidants, as in ozone pollution or the therapeutic use of pure oxygen; it also occurs when there is excessive accumulation of leucocytes\textsuperscript{29} as in the lungs of the heavy smoker.

Why has evolution provided such a vulnerable reactive centre for α1-antitrypsin? The answer seems to be that the ability to switch off α1-antitrypsin provides a means by which essential tissue breakdown can occur in the immediate vicinity of an inflammatory focus\textsuperscript{3} and perhaps in areas of regrowth. As an example, consider the tissue reaction to a thorn in the thumb. The thorn is rapidly surrounded by leucocytes which must: (1) kill any associated bacteria by discharging oxygen free radicals, and (2) isolate the focus by releasing proteases to give tissue liquefaction. This liquefaction can occur in the immediate vicinity of the inflammatory focus because the oxygen free radicals released by the leucocytes inactivate α1-antitrypsin by oxidizing its reactive centre methionine. The oxygen free-radicals have only a short radius of activity and beyond this area the unoxidized α1-antitrypsin readily neutralizes the elastase. In this way the vulnerability of the methionine becomes advantageous by allowing a defined zone of elastase activity.

However, evolution has obviously not had time to adjust to the introduction of cigarette smoking. Here the ability of oxidants to switch off the inhibitory activity of α1-antitrypsin becomes grossly disadvantageous because the smoker has an accumulation of leucocytes in the lungs and hence a combined exposure to elastase release and oxidative inactivation of α1-antitrypsin.

The Pittsburgh variant and site-directed mutation

A theme of this review is that the specificity of the serpins is primarily dependent on a single amino acid at their reactive centre. The possibility that this specificity of inhibition could be changed by altering the reactive centre amino acid became a reality with the identification of an extraordinary natural mutation. This was found in a 14-year-old male from Pittsburgh with a severe bleeding disorder caused by greatly increased antithrombin activity.\textsuperscript{7} The presence of an electrophoretically abnormal α1-antitrypsin suggested that the disease was caused by an active-site mutation. The results of analysis confirmed this in a most satisfying way.\textsuperscript{28}

As predicted, there was a mutation of the reactive centre methionine 358 of α1-antitrypsin to arginine (Table I). In effect, the inhibitor had been converted from a Meterserin to an Argserpin, i.e. it lost the ability to inhibit elastase and became a highly effective inhibitor of thrombin (Fig. 4).

The identification of α1-antitrypsin Pittsburgh raised the possibility of engineering mutations of the reactive centre. The natural mutation in the Pittsburgh variant has produced a potent anticoagulant with obvious therapeutic applications. This potential has recently been achieved by Courtney and colleagues\textsuperscript{8} who used site-specific mutagenesis to give high yields of the 358 arginine mutant in E. coli. By varying the replacement amino acid it should be equally possible to engineer specific inhibitors of the other serine proteinases of the complement and inflammatory cascades.

One immediate question\textsuperscript{1-3} was whether it would be possible to produce an oxidation-resistant α1-antitrypsin. Intravenous replacement therapy with plasma concentrates of α1-antitrypsin is already being tested on patients with the genetic deficiency. Meeting the worldwide needs for such therapy will require microbial production of α1-antitrypsin using recombinant DNA techniques. Why not get the microbes to produce an improved α1-antitrypsin?\textsuperscript{2}

The obvious way of doing this\textsuperscript{1-3} is to replace the reactive centre methionine by a valine. Good yields of the 358 valine-α1-antitrypsin (Valserpin) have been achieved in yeast by Hallewell and colleagues\textsuperscript{1}, and in E. coli by other

---

**Fig. 4.** Gelatine plate incorporating an inhibitor to give end-point measurements of porcine pancreatic elastase activity before and after oxidation with chloramine T. The control activity, without inhibitor, is shown in 1. Addition of both plasma (2) and yeast (4) α1-antitrypsin (Meterserin) inhibits elastase but only the valine mutant (Valserpin) of yeast (5) retains inhibitory activity after oxidation. The Pittsburgh mutant (Argserpin) is an ineffective elastase inhibitor (3) but an effective thrombin inhibitor \( K_{\text{inact}} = 1 \times 10^{-7} \) fully equivalent to heparin-activated antithrombin-III. Plate prepared by P. M. George.
Carbohydrate structures of glycoproteins and glycolipids as differentiation antigens, tumour-associated antigens and components of receptor systems

T. Feizi and R. A. Childs

Studies with monoclonal antibodies have shown that a family of blood group-related antigens based on the carbohydrate backbone sequences Galβ1-4(or3)GlcNAc behave as tumour-associated antigens in certain cell types of certain individuals but as normal antigens in others. Also, some of the rapid changes in antigenicity during embryogenesis result from structural changes in these saccharides. The saccharides are borne on glycoproteins and glycolipids and there is evidence that they are associated with receptor systems such as the receptor glycoprotein for the epidermal growth factor. Sialyl-oligosaccharide sequences in this series are host-cell receptors for the factor. Sialyl-oligosaccharide sequences in this series are host-cell receptors for the...