A novel pathological mutant reveals the role of torsional flexibility in the serpin breach in adoption of an aggregation-prone intermediate

Kamila Kamuda1,2, Riccardo Ronzoni1,2, Avik Majumdar3,4,5, Fiona H. X. Guan3, James A. Irving1,2, and David A. Lomas1,2

1 Division of Medicine, UCL Respiratory, Rayne Institute, University College London, UK
2 Institute of Structural and Molecular Biology, Birkbeck College, University College London, UK
3 AW Morrow Gastroenterology and Liver Centre, Royal Prince Alfred Hospital, Sydney, Australia
4 Victorian Liver Transplant Unit, Austin Health, Melbourne, Australia
5 The University of Melbourne, Melbourne, Australia

Keywords
alpha-1-antitrypsin; alpha-1-antitrypsin deficiency; conformational disease; protein folding; serpin

Correspondence
J. A. Irving, Division of Medicine, UCL Respiratory, Rayne Institute, University College London, London WC1E 6JF, UK
Tel: +44 2031087735
E-mail: j.irving@ucl.ac.uk

James A. Irving and David A. Lomas joint senior authors

(Received 24 November 2023, revised 17 February 2024, accepted 7 March 2024)
doi:10.1111/febs.17121

Introduction

Mutants of alpha-1-antitrypsin cause the protein to self-associate and form ordered aggregates (‘polymers’) that are retained within hepatocytes, resulting in a predisposition to the development of liver disease. The associated reduction in secretion, and for some mutants, impairment of function, leads to a failure to protect lung tissue against proteases released during the inflammatory response and an increased risk of emphysema. We report here a novel deficiency mutation (Gly192Cys), that we name the Sydney variant, identified in a patient in heterozygosity with the Z allele (Glu342Lys). Cellular analysis revealed that the novel variant was mostly retained as insoluble polymers within the endoplasmic reticulum. The basis for this behaviour was investigated using biophysical and structural techniques. The variant showed a 40% reduction in inhibitory activity and a reduced stability as assessed by thermal unfolding experiments. Polymerisation involves adoption of an aggregation-prone intermediate and paradoxically the energy barrier for transition to this state was increased by 16% for the Gly192Cys variant with respect to the wild-type protein. However, with activation to the intermediate state, polymerisation occurred at a 3.8-fold faster rate overall. X-ray crystallography provided two crystal structures of the Gly192Cys variant, revealing perturbation within the ‘breach’ region with Cys192 in two different orientations: in one structure it faces towards the hydrophobic core while in the second it is solvent-exposed. This orientational heterogeneity was confirmed by PEGylation. These data show the critical role of the torsional freedom imparted by Gly192 in inhibitory activity and stability against polymerisation.

Abbreviations
5k PEG, methoxypolyethylene glycol maleimide; AAT, alpha-1-antitrypsin; AATD, alpha-1-antitrypsin deficiency; CD, far-ultraviolet circular dichroism; COSM, centre of spectral mass; DTT, dithiothreitol; \(E_a\), apparent activation energy; ER, endoplasmic reticulum; FRET, fluorescence resonance energy transfer; IF, immunofluorescence; \(K_q\), quenching constant; \(M^\#\), polymerisation-prone intermediate; RCL, reactive centre loop; SD, standard deviation; SEM, standard error of the mean; \(T_m\), midpoint temperature of change.
resulting in accumulation of aggregates. These lead to conformational diseases which encompass a diverse group of disorders across all tissues of the body [1]. Among these, the serpinopathies are characterised by the formation of self-assembled, functionally inactive polymers of members of the serpin superfamily of proteinase inhibitors, predominantly accumulating within the lumen of the ER [2]. The chronically overloaded ER undergoes morphological distortion and dysfunction that is associated with disease.

Alpha-1-antitrypsin (AAT) is a 52 kDa, 394 amino acid glycoprotein that represents the archetypal member of the serpin superfamily. It is predominantly synthesised within the liver and then secreted into the circulation where it carries out its function as a proteinase inhibitor, mainly inhibiting neutrophil elastase [3]. The structure of AAT is based around a central $\beta$-sheet A and a 20 amino acid reactive centre loop (RCL) [4]. Many mutations of AAT have been described to destabilise the native protein allowing it to form ordered polymers that are retained as PAS-positive, diastase-resistant inclusions within the ER of hepatocytes [5]. The retention of these polymers is associated with cirrhosis and hepatocellular carcinoma while the reduced AAT secretion predisposes to early-onset emphysema; these conditions represent the characteristic pathology associated with alpha-1-antitrypsin deficiency (AATD). The best-characterised of the deficiency variants associated with AATD is the Z allele (Glu342Lys) [6], but different mutations including Siyama (Ser53Phe) [7], Mmalton ($\Delta 52$) [8] and King’s (His334Asp) [9] similarly form polymers that are retained within hepatocytes and predispose to liver disease. These and other naturally occurring mutations have played a fundamental role in advancing understanding of the structural determinants of stability, function and misfolding of serpins [10]. During inhibition of a target protease [11,12] and likely during polymerisation [6], a serpin undergoes a conformational rearrangement in which $\beta$-sheet A accommodates the RCL as a newly formed sixth strand. A key region involved in the initiation of this structural transition is the breach. The breach region is located in the vicinity of the top of the strands 3 and 5 of $\beta$-sheet A, as well as strands 2, 3 and 4 of $\beta$-sheet B [13] and serves as a primordial point of insertion of the RCL (Fig. 1).

We report here a novel mutant of AAT identified in an individual from Sydney, Australia. Characterisation of this variant in cellular models of disease and using structural studies showed that Cys192 perturbs the structure of the breach region of the protein to form the intracellular polymers that are associated with liver disease in AATD. The effect of this mutation highlights the dual role played by the breach in both inhibitory activity and polymerisation. Interestingly, in comparison with other mutants that we have characterised, this variant acts by destabilising the polymerisation intermediate.

Fig. 1. The native AAT structure. A canonical ‘front’ and ‘back’ view of AAT (PDB: 1QLP) [27] highlighting the functionally relevant breach region, which initiates insertion of the reaction centre loop (RCL) into $\beta$-sheet A. Structures were generated using CHIMERA [56].
Results

Clinical details

A 32-year-old male was identified at the Royal Prince Alfred Hospital in Sydney with abnormal liver enzymes; ALT, AST and gamma GT were elevated at 67 U·L\(^{-1}\), 43 U·L\(^{-1}\) and 50 U·L\(^{-1}\) respectively. Alkaline phosphatase and serum bilirubin were in the normal range and synthetic liver function was intact. There was no evidence of autoimmune and viral causes of liver disease. Liver stiffness measurement with Fibroscan\textsuperscript{®} (Echosens, Paris, France) was 5 kPa in the non-fasting state with a controlled attenuation parameter score (CAP) of 260. AAT levels were low in a range between 0.36 and 0.54 g·L\(^{-1}\) (normal range 1.5–3.5 g·L\(^{-1}\)), approximately in the range seen with individuals with a ZZ genotype [14]. Phenotyping revealed the presence of the Z allele (Glu342Lys) in heterozygosity with a second, unknown allele. Genotyping of the AAT (SERPINA1) gene identified a novel mutation c.646G>T transversion that results in the substitution of glycine to cysteine at position 192 (Gly192Cys). This new protein variant was named AAT Sydney (here referred to as Gly192Cys) after the name of the city in which it was identified. The circulating concentration of AAT was below that seen with SZ heterozygotes and so this variant could be considered as severely deficient [15]. Lung function tests on this patient showed an FEV\(_1\) of 3.21 L (76% of predicted) and an FVC of 4.28 L (83% predicted; ratio 75% predicted). The FEV\(_1\) increased to 3.83 L after 400 µg of inhaled salbutamol (19% increase). Static lung volumes were normal and the corrected K\textsubscript{co} was 111% of predicted.

Gly192Cys AAT accumulates as polymers within the cell

To evaluate the basis for the deficiency seen in the proband, Hepa1.6 cells were transiently transfected with Gly192Cys AAT or with the wild-type M or the Z mutant (Glu342Lys) for reference. The variants were expressed for 20 h and the cells were lysed in 1% v/v NP-40 detergent in order to resolve the soluble and insoluble intracellular fractions as previously described [16], while cell media was collected to investigate the extracellular secreted fraction. The Gly192Cys mutation introduces a second cysteine in addition to that in the wild-type protein (Cys232) and so could promote the formation of intra- and intermolecular disulphide bonds during folding; this has previously been observed for the I and Brixia variants (Arg39Cys and Phe35Cys, respectively) [17]. The cellular accumulation of Gly192Cys AAT was investigated by resolving both the intracellular and the extracellular components in the presence or absence of reducing agent (DTT). Under non-reducing conditions, Gly192Cys AAT formed a high molecular weight complex, present in the soluble and insoluble cellular fractions (Fig. 2A, top panel), which was also observed in Z AAT. Additionally, Gly192Cys AAT formed an intrachain disulphide bond (Fig. 2A, marked with an asterisk), which migrated faster than the 52 kDa intracellular monomeric form. This band was present both in the soluble and insoluble fraction and was indeed similar to that seen with the I and Brixia cysteine mutants [17]. These intra- and interchain complexes dissociated in the presence of DTT (Fig. 2A, middle panels), confirming they are linked by disulphide bonds. The accumulation of Gly192Cys AAT in the insoluble fraction suggests that this variant forms intracellular polymers similar to those of Z AAT. To assess this, an immunoprecipitation with the anti-AAT polymer 2C1 monoclonal antibody [9] was performed (Fig. 2A, bottom panel). The relative levels of the mature glycosylated form in the extracellular fraction indicated a more efficient secretion of Gly192Cys AAT than Z AAT (Fig. 2B). Cellular immunofluorescence showed a characteristic puncta-like pattern for Gly192Cys AAT, which is consistent with polymers of AAT forming inclusion bodies as observed for the cells expressing Z AAT (Fig. 2C) [9,18,19]. This pattern was not seen in cells expressing M AAT, where a majority of the protein signal accumulates in the perinuclear area, consistent with protein processing through the Golgi apparatus. Interestingly, the perinuclear staining pattern was not observed in cells expressing Gly192Cys AAT, which might suggest an increased rate of degradation or a slower rate of protein secretion from the cells.

The Gly192Cys mutation compromises functional activity

Purified recombinant Gly192Cys AAT (Fig. 3A, inset) was subjected to a series of biochemical assays in comparison with recombinant M AAT. The Z variant, which destabilises the breach, leads to an ~30% reduction in inhibitory activity [20]. The Gly192Cys mutation had a similar effect with a 1.7-fold molar excess being required to achieve the complete inhibition of the model protease α-chymotrypsin (Fig. 3B) [21].

Fluorescence spectroscopy suggests exposure of a breach region

A series of spectroscopic experiments were used to investigate the bulk solution-state properties of Gly192Cys AAT.
AAT in comparison to the wild-type (M) protein. Far-ultraviolet circular dichroism spectroscopy (CD) (Fig. 3A) revealed a near-identical profile for M and Gly192Cys AAT showing that the mutation did not perturb the global structure of the protein. An evaluation of its intrinsic fluorescence profile, which has been shown to primarily report changes in the environment around Trp194 situated within the breach region (Fig. 1), was undertaken. In common with AAT Z [22], the emission profile showed increased fluorescence intensity for Gly192Cys AAT relative to the wild-type protein (Fig. 3C). However, in contrast to Z AAT, there was no observable shift in the centre of spectral mass (COSM) relative to the wild-type. This suggests little change of polarity in the immediate environment of Trp194, while the increased fluorescence quantum yield likely reports other rearrangements in the local environment that favour radiative processes over non-radiative dissipation of energy.
Bis-ANS is an environmental reporter dye that has been shown to selectively bind Z AAT with respect to the wild-type protein at moderate temperatures resulting in a pronounced increase in its fluorescence quantum yield [20,22,23]. Similarly, in the presence of the Gly192Cys AAT mutant at 25°C, bis-ANS...
exhibited an increased emission intensity (Fig. 3D). In combination, these results suggest similarities in the perturbation of the molecule induced by the novel mutation, with a lesser impact on its global structure.

**Gly192Cys destabilises the intermediate relative to the native state**

A thermal stability assay reports the midpoint ($T_m$) at which 50% of the protein unfolds. In the case of AAT, this ‘unfolded’ state represents the transition to an intermediate ensemble on the polymerisation pathway [24]. The dependence of the $T_m$ on the rate of heating can allow calculation of the apparent activation energy of the reported transition [25]. M and Gly192Cys AAT were heated from 20 to 94 °C at four rates – 0.5, 1, 2 and 4 °C-min$^{-1}$ (Fig. 4A). At each ramp rate, the observed $T_m$ for Gly192Cys AAT was lower than that of the wild-type protein (Table 1). Interestingly, the height of the apparent energy barrier ($E_a$) for the intermediate state, calculated from the regression slopes of $T_m$ at different ramp rates

![Graphs and images illustrating the thermal stability assay and data analysis](image-url)
The Gly192Cys mutation destabilises the native and intermediate AAT conformations. (A) The progress curves of thermal shift experiments in which the transition of the monomeric protein was monitored in the presence of SYPRO Orange from 25 to 90 °C at four different rates of temperature increase (°C-min⁻¹). (B) The midpoint of the thermal transition (Tm) was plotted as a function of ramp rate, according to the numerical transformations shown. (C) The histograms of calculated apparent activation energy values (Ea) for the transition from native to intermediate states were obtained from the slopes of the regression curves shown in (B). The data are representative of the average of triplicates in two independent experiments, while error bars indicate standard deviation (n=2, ± SD). M and Gly192Cys AAT are shown in blue and pink, respectively. (D) The polymerisation of AAT variants labelled with Atto-488-NHS and Atto-594-NHS dyes was monitored over a range of temperatures from their increase in relative FRET signal. The polymerisation progress curves were used to determine half-times of polymerisation (t½). (E) The apparent activation energy for polymerisation (Ea) was calculated from an Arrhenius plot using these values. (F) The half-time (t½) of polymerisation at a reference temperature (54.8 °C) to provide a typical single-temperature comparison was interpolated from the Arrhenius plot from the FRET polymerisation assay. (G) A representative non-denaturing PAGE for the heat-induced polymerisation assay in which 0.2 mg mL⁻¹ AAT was incubated for 4 h at various temperatures with separation using a 3–12% w/v acrylamide non-denaturing electrophoresis. Plasma M (pM) and Z (pZ) AAT were used as controls to distinguish the migration pattern of the recombinant proteins. RT represents incubation at room temperature. The statistical significance in (E) and (F) was assessed by unpaired student t-test; **** represents P < 0.0001. The data shown in D–F are representative of three separate experiments, error bars indicate standard error mean (n=3, ± SEM).

**Table 1.** Observed thermal midpoints (Tm) for M and Gly192Cys AAT at four different ramp rates (n=2, ± SD).

<table>
<thead>
<tr>
<th>Rate (°C-min⁻¹)</th>
<th>Tm (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>M: 56.5 ± 0.3, Gly192Cys: 52.7 ± 0.2</td>
</tr>
<tr>
<td>1</td>
<td>M: 58.6 ± 0.1, Gly192Cys: 54.5 ± 0.4</td>
</tr>
<tr>
<td>2</td>
<td>M: 61.4 ± 0.1, Gly192Cys: 57.4 ± 0.3</td>
</tr>
<tr>
<td>4</td>
<td>M: 63.0 ± 0.1, Gly192Cys: 58.0 ± 0.5</td>
</tr>
</tbody>
</table>

(Fig. 4B), was higher for Gly192Cys AAT in comparison with M AAT (326.6 kJ mol⁻¹ ± 33.1 and 282.2 kJ mol⁻¹ ± 12.8, respectively) (Fig. 4C).

The Gly192Cys mutant leads to an increase in the rate of polymerisation

Polymerisation can be induced artificially under destabilising conditions; thermal stress results in the adoption of a polymer form that shares structural characteristics and exposed epitopes with those that arise naturally in the liver [9,26]. The kinetics of polymerisation were followed by FRET using AAT labelled with two fluorophores acting as a fluorescence donor and acceptor [20,21]. The half-time of FRET increase at temperatures between 50 and 65 °C was represented on an Arrhenius plot as a function of temperature (Fig. 4D). The slopes of the regression lines indicate the energy barrier for polymerisation to progress, which was found to be reduced for the Gly192Cys mutant (Fig. 4E). The shorter half-times for Gly192Cys AAT are reflected in the shift of the regression line and suggest a faster polymerisation rate. The regression curves also allowed a more accurate comparison, by interpolating from the rates at all temperatures, of the half-time for polymerisation at a single temperature (Fig. 4F), which at 54.8 °C was 3.8 times faster for Gly192Cys than M AAT.

Crystalllographic analysis

Crystallisation of Gly192Cys AAT using the hanging drop vapour diffusion method resulted in the formation of plate-like crystals, which were screened at the Diamond Light Source synchrotron (Harwell Science and Innovation Campus, Didcot, UK). Using the highest resolution dataset, in the C2 space group, the mutant structure was solved at a resolution of 1.9 Å (PDB: 8P4J) (Table 2). A second crystal structure of Gly192Cys mutant (PDB: 8P4U) was also solved at 2.4 Å resolution (Table 2).

In the native structure of Gly192Cys AAT (PDB: 8P4J; Table 2), the additional Cys192, located at the top of strand s3A of β-sheet A, has a side chain oriented towards the inner face of the protein (Fig. 5A). Interestingly, the cysteine is within the breach, near the disease-associated Z mutation which is located at the head of strand s5A (Fig. 5A, right panel).

The effect of the mutation was determined by comparison with our crystal structure of M AAT (PDB: 1QLP) [27] based on a backbone alignment of the two molecules (Fig. 5B). The substitution of glycine with
cysteine caused structural rearrangements within the breach region. In M AAT, the side chain of Lys343 is normally solvent-exposed. However, in the Gly192Cys mutant, Lys343 leans towards the loop connecting strand 3A and β-sheet C, creating a new hydrogen bond with the backbone of Trp194 at a distance of 3.2 Å. Additionally, the presence of the cysteine side chain in the previously unoccupied space (due to the lack of a side chain in glycine) between β-sheet A and β-sheet B, resulted in rearrangement of Tyr244. As a consequence, the tyrosine ring is pushed back towards β-sheet B causing small movements of strands s3B and s4B (Fig. 5C). Tyr244 is conserved in over 70% of serpins [11], and its orientation as found in wild-type AAT is conserved in over 90 PDB structures of serpins in various conformations deposited in the RSCB Protein Data Bank. This finding makes this mutant structure unique among known mutants of AAT.

Further analysis was undertaken to compare the B-factors for two M and two Gly192Cys AAT structures (PDB: 1QLP, PDB: 3NE4, PDB: 8P4J and PDB: 8P4U respectively). Increased B-factor values were evident in the region of the mutation and in the loop connecting strands 3 and 4 of β-sheet C (Fig. 5C), suggesting higher flexibility or disorder in the mutant. Moreover, slight variations were noticed around α-helix G and H as well as in the RCL. It is important to note that the RCL is highly flexible due to its protease inhibitory function, as it serves as an exposed docking site for the protease. Consequently, this region is often disordered in AAT structures, and in the case of the structure of Gly192Cys AAT (PDB: 8P4J), eight RCL residues (344–351) lacked sufficient electron density to be modelled.

The differences in normalised B-factors between Gly192Cys (PDB: 8P4J) and M (PDB: 1QLP) AAT were mapped onto the 8P4J structure to highlight the distribution of differences on the tertiary protein structure (Fig. 5D). The absolute difference in B-factors revealed the majority of the protein structure to be unchanged or only moderately perturbed. However, more marked differences were observed around the breach region at the site of the mutation along the top of β-sheet A, suggestive of intrinsic local disorder introduced by Cys192. There was no evidence of a distance-dependent propagation of increased B-factors through the molecule (data not shown), which has been observed in NMR chemical shift perturbations in other proteins [28]. The underlying region of strands 4 and 5 of β-sheet B were also affected which can be explained by the accommodation of Cys192 in the inner cavity within the breach. Interestingly, a B-factor difference was observed in the solvent-exposed strands 3 and 4 of β-sheet C as well as α-helix H and a loop

### Table 2. Data collection and refinement statistics*.

<table>
<thead>
<tr>
<th>Data collection</th>
<th>8P4J</th>
<th>8P4U</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wavelength, Å</td>
<td>0.9795</td>
<td>0.9703</td>
</tr>
<tr>
<td>Space group</td>
<td>C 1 2 1</td>
<td>C 1 2 1</td>
</tr>
<tr>
<td>Cell dimensions</td>
<td>a, b, c (Å)</td>
<td>114.1, 38.5, 90.0</td>
</tr>
<tr>
<td></td>
<td>α, β, γ (°)</td>
<td>90, 104.6, 90</td>
</tr>
<tr>
<td>Resolution range (Å)</td>
<td>87.08–1.98 (95.1–1.98)</td>
<td>42.67–2.20 (2.56–2.20)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>64.9 (12.4)</td>
<td>61.4 (12.1)</td>
</tr>
<tr>
<td>Completeness</td>
<td>84.1 (34.2)</td>
<td>87.5 (50.9)</td>
</tr>
<tr>
<td>Refinement</td>
<td>Rwork/Rfree</td>
<td>0.21 (0.28)</td>
</tr>
<tr>
<td>No. reflections</td>
<td>19,263 (79)</td>
<td>11,027 (298)</td>
</tr>
<tr>
<td>No. atoms</td>
<td>2767</td>
<td>2697</td>
</tr>
<tr>
<td>R-factors</td>
<td>31.80</td>
<td>50.18</td>
</tr>
<tr>
<td>R.m.s. deviations</td>
<td>Bond lengths (Å)</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>Bond angles (°)</td>
<td>0.560</td>
</tr>
<tr>
<td>Ramachandran plot</td>
<td>Favored, %</td>
<td>95.71</td>
</tr>
<tr>
<td></td>
<td>Allowed, %</td>
<td>4.29</td>
</tr>
<tr>
<td></td>
<td>Disallowed, %</td>
<td>0.0</td>
</tr>
</tbody>
</table>

*Brackets indicate values for the reported high resolution range.
A breach region glycine and serpin polymerisation

(A) 

(B) 

(C) 

(D) 

(E)
connecting strands 4 and 5 of β-sheet B, which are not in direct proximity to the site of the mutation.

**Two conformational forms of Gly192Cys AAT**

Only the thiol head of Cys192 could be placed within the electron density of the lower resolution crystallographic dataset (PDB: 8P4U) (Fig. 6A). Electron density was absent from the inner face of the breach region and Tyr244 was shifted back into the breach region pocket as in native AAT, and in contrast to the initial Gly192Cys dataset (PDB: 8P4J). The difference in the shift of Cys192 and Tyr244 between the higher and lower resolution structures was

---

(C) | M | G192C
---|---|---
DTT | + | - | + | -
50 kDa | + | 10 kDa | + | 5 kDa
AAT | | | | |

---

(D) | [Acrylamide] (M) | 0.0 | 0.1 | 0.2 | 0.3 | 0.4 | 0.5
---|---|---|---|---|---|---|---
F_0/F | | | | | | | |
M | | | | | | | |
G192C | | | | | | | |

(E) | KSV (M⁻¹) | 2 | 4 | 6 | 8 | 10
---|---|---|---|---|---
M | | | | | |
G192C | | | | | |
Fig. 6. The orientation of cysteine 192 at the mutation site. (A) Placement of Cys192 in the 2Fo–Fc electron density maps for two solved structures: 8P4J in pink and 8P4U in dark pink, contoured at 1σ. (B) Structural comparison of the mutation site from two Gly192Cys AAT crystal structures (PDB: 8P4J and PDB: 8P4U). The top panel shows AAT residues 190–194 and 243–249 and their respective 2Fo–Fc electron densities, contoured at 1σ. The bottom panels show superimposed structures, with their electron densities (left) and in the ribbon representation with highlighted Cys192 and Tyr244, and the distances between the residues in the two structures (right). (C) A representative 4–12% w/v acrylamide SDS/PAGE gel from the PEGylation experiments (n = 3), in which covalent binding of 5k PEG to one cysteine results in a molecular weight shift of 5kDa. As a control, prior to labelling, protein variants were also reduced with DTT. (D) Stern-Volmer plots for fluorescent quenching with acrylamide. AAT variants were at 5μM (blue and pink, respectively for M and Gly192Cys AAT). (E) The quenching constant (KSV) values, derived from the slope of the regression line in panel (D). The values reported in panels D and E are the averages from three independent experiments, error bars indicate standard deviation (n = 3, ± SD).

5.7 Å (Fig. 6B). This suggests that Gly192Cys presents as two conformers, that possibly alternate between one another.

To investigate the apparent rotational flexibility evident at position 192 in the two crystal structures, further biochemical assays were performed to assess the orientation of Cys192. Firstly, M and Gly192Cys AAT were subjected to conjugation with 5 kDa methoxypolyethylene glycol maleimide (5k PEG) which covalently binds free cysteines. AAT has one cysteine at position 232, at the top of strand 1 of β-sheet B and is accessible for labelling in this way [29], which resulted in a ~5 kDa molecular weight shift (Fig. 6C). The conformation of the second cysteine might be determined by either a single or double shift in molecular weight (~5 or ~10 kDa, respectively). Two labelled populations of the mutant were seen on SDS/PAGE, one with an ~5 kDa and one with ~10 kDa molecular weight shift (Fig. 6C). This suggests that Gly192Cys AAT is present in two protein conformations, in one of which Cys192 is buried and not accessible for conjugation, and a second in which the mutated residue is solvent exposed.

As described above, it has been found that Trp194 is the primary contributor to the intrinsic fluorescence signal of AAT [30,31], which sits within the breach region. The effect of the Gly192Cys mutation on the solvent accessibility to Trp194 was assessed in a fluorescent quenching assay in which the tryptophan signal was suppressed by stepwise addition of acrylamide, with the fluorescence intensity plotted as a function of acrylamide concentration (Fig. 6D). The quenching constants (KSV) derived from the slope of the regression (Fig. 6E) were: 6.1 ± 1.0 M⁻¹ (SEM) and 7.3 ± 1.1 M⁻¹ (SEM) for M and Gly192Cys AAT, respectively. A higher KSV is associated with a faster quenching rate, which indicates increased solvent exposure of the breach region, possibly caused by alternating conformations of cysteine. This may reflect the buried residue reducing quencher access to Trp194 whereas the solvent-facing cysteine exposes the breach region of the protein. The curvature on the Stern-Volmer plot was consistent with a “quenching sphere of action” in which proximity of the quencher to some of the fluorophores immediately dissipates excitation energy [32] (Fig. 6D).

Discussion

The novel Gly192Cys mutation was identified in an individual with low circulating levels of AAT. The site of the substitution is in the breach region which is involved with both inhibitory activity and polymerisation. Formation of the AAT:protease complex requires a rapid conformational transition which results in the opening of β-sheet A and insertion of the RCL as strand 4A [33,34] initiated within the breach region [13]. The importance of this for the serpin superfamily is reflected in the generally high sequence conservation within the breach [11]. Wild-type AAT sequence alignments with human inhibitory serpins and sequences of AAT from different species support a highly conserved glycine at position 192 (Fig. 7A,B). Among amino acids, glycine exhibits the greatest degree of rotational freedom due to the lack of a side chain. Conservation of this residue at the top of strand 3A therefore suggests a requirement for torsional flexibility for the proper functioning of the breach, as the protein goes through a series of intermediate states to achieve the thermodynamically favourable loop-inserted form [35]. Consistently, we observed an ~40% decrease in inhibitory activity of the AAT Gly192Cys variant, whose cysteine residue imposes a greater torsional restriction (Fig. 3B). In a comparison of 21 inhibitory human serpins, most also had a glycine residue, although four instead tolerated an alanine at this position. A superimposition of crystal structures indicated that these serpins have adapted to this residue by adopting a common side chain orientation away from the breach, which differs from that seen in the variant structures reported here (Fig. 7C).

The FEBS Journal (2024) © 2024 The Authors. The FEBS Journal published by John Wiley & Sons Ltd on behalf of Federation of European Biochemical Societies.
These crystallographic studies identified two orientations of Cys192 within Gly192Cys AAT (Fig. 6). In the 8P4J structure, Cys192 faces towards the inner face of the protein, leading to notable rearrangements, especially of the highly conserved Tyr244 residue. Interestingly, in the 8P4U structure, where the cysteine is solvent-exposed, Tyr244 is repositioned to the canonical orientation (Fig. 6). The presence of a mixed monomeric population in solution was confirmed by the PEGylation assay with a visible double shift in molecular weight by the conjugation of both single and double adducts (Fig. 6C), and supported by the acrylamide...
tion was destabilised, the relative activation energy ($E_a$) of this change [25]. Paradoxically, while the native configuration was destabilised, the relative $E_a$ barrier for formation of the intermediate was higher for Gly192Cys in comparison with M AAT (Fig. 4B,C). The structure of the polymerisation intermediate can be inferred from a small molecule-bound form of AAT [37,38]; interestingly, in this form the backbone is perturbed in the vicinity of position 192, relative to the native state (Fig. 7D). Interchange between the buried and exposed orientations of the cysteine would be reduced in comparison with the torsional freedom of a glycine residue, partially impeding intermediate formation, but favouring polymer formation over reversion to the native conformation once it has formed. Indeed, the thermal stability experiments indicated a lower degree of reversibility to the native state.

A further possible contributor to the increased activation barrier is the observed structural perturbation of Lys343 (Fig. 5B, inset panel 1) which involves the formation of a new stabilising hydrogen bond. Further biochemical analysis indicated a difference in the interaction of this molecule with the environmental reporter bis-ANS [22], which has been found to recognise an intermediate conformation of AAT that is populated by the Z variant of AAT and by wild-type M AAT when it is heated [39]. Similar to Z AAT, Gly192Cys AAT induced a much higher fluorescence intensity in the dye over the wild-type protein (Fig. 3D). While the site of interaction of this molecule has not been established, the observed increased movements of hH and s3B and s4B suggests that destabilisation of $\beta$-sheet B (Fig. 5D) may be involved in this process.

Once activated into the polymerisation-prone intermediate, the energy barrier for the irreversible self-association of the mutant with another molecule is significantly reduced, leading to a faster rate of polymerisation upon thermal treatment in comparison to the wild-type protein (Fig. 4). When expressed in Hepa1.6 cells, Gly192Cys AAT shared with Z AAT the ability to form intracellular polymers and inclusion bodies within the ER of cells (Fig. 2). This accumulation has also been observed in other rare variants such the highly polymerogenic Baghdad (Ala336Pro) [20] and King’s (His334Asp) [9] AAT in which the functional opening of $\beta$-sheet A is impaired. It is also possible that Gly192Cys and Z AAT may interact and copolymerise during expression, by analogy with other variants [40,41].

The impact of the mutation on the severity of the liver and lung disease is hard to characterise clinically as the index case is also a carrier of the severely defective Z allele. Nevertheless, experiments performed in a mammalian expression system, as well as biochemical, biophysical and crystallographic studies, collectively demonstrate that the Gly192Cys mutation promotes formation of AAT polymers associated with an increased risk of liver disease, and depletion of functional levels associated with lung disease. However, this is likely to be to a lesser degree than that seen with homozygosity for the severe Z deficiency allele.

### Materials and methods

Unless otherwise noted reagents were obtained from Sigma Aldrich (St. Louis, MO, USA)/Merck (Darmstadt, Germany).

### Cloning and plasmid preparation

The QuickChange II site-directed mutagenesis kit (Agilent, Santa Clara, CA, USA) was used to introduce the single point mutation (Gly192Cys) into wild-type AAT encoded by recombinant and mammalian vectors. Recombinant proteins were expressed using the pQE-30 vector (Qiagen, Hilden, Germany) containing the AAT sequence with an N-terminal hexa-histidine tag for affinity chromatography [24]. The expression vectors used for the mammalian cell culture assays were as previously reported [42] and based on those encoding human M1 (Val213) and Z AAT.

### Recombinant protein expression and purification

Recombinant AAT was expressed in XL1-Blue E. coli cells (Agilent) and purified by affinity chromatography as described previously [24] with the addition of 100 mM $\beta$-mercaptoethanol prior to ion-exchange. The yield of monomeric Gly192Cys AAT was $\sim 8 \text{ mg L}^{-1}$ of culture, approximately 50% lower than that of M AAT under equivalent conditions. Protein homogeneity was established by non-denaturing and SDS/PAGE electrophoresis. The resulting purified protein was buffer-exchanged into either PBS (PBS pH 7.4, 0.02% w/v Na$_3$PO$_4$) or crystallography buffer (10 mM Tris pH 7.4, 50 mM NaCl, 0.02% w/v Na$_3$PO$_4$) and stored at $-80^\circ$C.
Mammalian cell culture

The Hepa1.6 mouse hepatoma cell line (ATCC CRL-1830) was grown in Dulbecco’s modified eagle medium (DMEM) (Sigma Aldrich) supplemented with 10% v/v foetal bovine serum. Transient transfections were performed with polyethyleneimine ‘MAX’ (PEI) (Polysciences, Warrington, PA, USA) as previously described [17]. Transfected cells were then washed twice with pre-warmed PBS and further incubated at 37 °C with 1 mL of Opti-MEM (Gibco, Thermo Fisher Scientific, Waltham, MA, USA). After 20 h incubation, cell culture media were collected, centrifuged at 800 g for 5 min at 4 °C, transferred into clean tubes and stored at −20 °C, while the cells were lysed as previously described [16].

Hepa1.6 cells, extensively utilised for AAT studies [16,43–45] had been procured from and authenticated by Merck KGaA (Darmstadt, Germany) < 3 years before the first transfection. All experiments were conducted using cells routinely screened for mycoplasma contamination.

Cellular immunofluorescence

As previously described in Ronzoni et al. [44], Hepa 1.6 cells were seeded onto 2 cm² coverslips (Millipore Sigma, Burlington, MA, USA) and transfected as described above. After 48 h, the cells were fixed with 4% v/v paraformaldehyde, permeabilised with 0.1% v/v Triton X-100, and immunodecorated with anti-human AAT (Agilent Dako, Glostrup, Denmark) (at a concentration of 2.2 μg·mL⁻¹) or the anti-AAT polymer 2C1 mAb [9] (0.8 μg·mL⁻¹) overnight at 4 °C. The primary antibodies were detected with goat anti-mouse antibody conjugated to Alexa Fluor 488 and goat anti-rabbit antibody conjugated to Alexa Fluor 555 (Thermo Fisher Scientific). Cells were counterstained with Hoechst (Thermo Fisher Scientific). The primary antibodies were detected with goat anti-mouse antibody conjugated to Alexa Fluor 488 and goat anti-rabbit antibody conjugated to Alexa Fluor 555 (Thermo Fisher Scientific) to visualise the nuclei. Coverslips were mounted on slides with Immuno-Mount (Thermo Fisher Scientific) and analysed on a Zeiss Airscan 880 confocal microscope with a 63x objective (1.4 oil).

Biochemical assays

The stoichiometry of inhibition (SI) of AAT against α-chymotrypsin was determined as described previously [21]. Intrinsic fluorescence spectra were recorded on a SpectraMax M5 plate reader (Molecular Devices, San Jose, CA, USA), whereas the bis-ANS readout was undertaken using a Cary Eclipse fluorescence spectrophotometer (Agilent Technologies), both at 25 °C [20]. For intrinsic fluorescence, samples were excited at 270 nm and scanned in the emission range from 290 to 450 nm. Bis-ANS was incubated with AAT for 10 min at room temperature prior to recording an emission scan between 400 and 700 nm with 10 nm slits for both emission and excitation and an excitation wavelength of 370 nm. Circular dichroism (CD) spectra were recorded at 25 °C on J-720 spectropolarimeter (Jasco, Inc., Easton, MD, USA) in the ISMB Protein Crystallography and Biophysics Centre (BiophysX). Protein samples were prepared in a low ionic strength buffer as described [21]. The stability of the AAT variants was determined in a SYPRO Orange-based thermal denaturation assay as described previously [20,21] (Thermo Fisher Scientific). The midpoint of denaturation (Tm) was determined by fitting the data to a nonlinear regression algorithm describing a two-state unfolding transition [24]. The kinetics of polymerisation for AAT labelled with Atto-488-NHS and Atto-594-NHS (Atto-Tec, Siegen, Germany) via primary amines were assessed in a FRET-based assay [20,21,24] using a RealPlex4 MasterCycler qPCR instrument (Eppendorf, Hamburg, Germany). Polymerisation was determined as an increase in FRET recorded at an emission wavelength 605 ± 15 nm upon excitation at 470 nm over a range of temperatures.

Crystallography and structure determination

Crystals of Gly192Cys AAT were grown in 100 mM of MIB buffer (sodium malonate dibasic monohydrate, imidazole, boric acid) (Molecular Dimensions, Sheffield, UK) and 25% v/v 1500 polyethylene glycol (Sigma). The protein used for crystallography was buffer-exchanged into crystallography buffer (10 mM Tris pH 7.4, 50 mM NaCl, 0.02% NaN3) and concentrated to 11 mg·mL⁻¹. Crystals were grown by the hanging drop technique when mixed in a 1 : 1 ratio with the growth buffer.

Fully-formed crystals (usually 7–10 days post-seeding) were snap-frozen in liquid nitrogen in the presence of cryoprotectant. Data collection was undertaken at the Diamond Light Source on the I03 beamline (Harwell Science and Innovation Campus). The diffraction data were recorded over 360° with 0.2° oscillation/frame. Collection conditions were as follows: beam wavelength 0.9763 Å, beam size 80 x 20 μm and average photon flux 1.95 x 10¹² and the exposure time was 0.05 s.

The collected diffraction data were autoprocessed with autoPROC and STARANISO [46,47] which performs an elliptical truncation of data and relies on other software, including XDS for space group determination and integration of diffraction data [48] and Aimless for merging and scaling [49].

The crystal structures were solved by molecular replacement with the native AAT structure [27] (PDB: 1QLP) used as an initial search model in Phaser [50], the atomic protein models were refined using PHENIX [51], and model building was undertaken with the COOT software [52]. The final crystal structures were validated using the online wwPDB validation server [53].

Polyethylene glycol conjugation assay

Conjugation of 5 kDa methoxypolyethylene glycol maleimide (5k PEG) (Sigma Aldrich) was undertaken by incubating the AAT variants at 0.5 mg·mL⁻¹ (11.1 μM) with a
Acrylamide-induced protein quenching assay

The quenching of tryptophan fluorescence was induced by acrylamide as recorded on a Cary Eclipse fluorescence spectrophotometer (Agilent Technologies). The wavelengths for excitation and emission were 280 and 340 nm, respectively. Five micromolar of AAT was prepared in a total volume of 90.5 μL and 0.1 μL of acrylamide was added five times stepwise upon the stabilisation of the fluorescence signal, to a final reaction volume of 100 μL.

The average of the raw fluorescent signal at each concentration of acrylamide \( F \) was normalised to the protein signal without acrylamide \( F_0 \) and corrected for the changing reaction volume. The normalised fluorescent signal was plotted as a function of the acrylamide concentration. Data were plotted using a correction factor for the ‘quenching sphere of action’ effect \([32]\) as follows:

\[
\frac{F_0}{F} = \left( \frac{Y_{\text{int}} + K_{\text{SV}}[Q]}{Y_{\text{int}}} \right) \ast e^{(V/Y)}
\]

where \( K_{\text{SV}} \) is the Stern-Volmer quenching constant, \([Q]\) is the quencher concentration, \( Y_{\text{int}} \) is the intercept of the line \((\text{around} \ 1)\) and \( V \) is related to a volume in which nearby quenchers entirely suppress fluorescence.

Sequence analysis

WebLogo \([54,55]\) was used to analyse and represent patterns of conservation. Sequences of antitrypsin orthologues were obtained from UniProt with the following accessions: P01009, A0A2K5DME6, A0A6J3ILL6, A0A2K6TI38, A0A2K7SNR2, A0A5E4D9W6, A0A2K5XHH4, A0A212CS37, A0A4X2KQL0, M3XVV7, A0A2Y9JIB0, Q00896, A0A5F4ZV20, O54761, H0WJ03, G3QXZ8, W5PZS7, A0A5F4ZV20, O54761, H0WJ03, G3QXZ8, W5PZS7, A0A5F4ZV20, O54761, H0WJ03, G3QXZ8, W5PZS7.

References

A breach region glycine and serpin polymerisation

K. Kamuda et al.


25. Faull SV, Elliston ELK, Gooptu B, Jagger AM, Aldobiani Y, Redzej A, Badoui M, Heyer-Chauhan N,


35 Stratikos E & Gettins PG (1999) Formation of the covalent serpin-proteinase complex involves translocation of the proteinase by more than 70 A and full insertion of the reactive center loop into beta-sheet A. *Proc Natl Acad Sci USA* 96, 4808–4813.


A breach region glycine and serpin polymerisation

K. Kamuda et al.