# Comprehensive clinical diagnostic pipelines reveal new variants in alpha-1-antitrypsin deficiency

Stefania Ottaviani[1], Giulia Bartoli[2], Tomás P. Carroll[3], Fabrizio Gangemi[2], Alice M. Balderacchi[1], Valentina Barzon[4], Alessandra Corino[1], Davide Piloni[1], Noel G. McElvaney[3], Angelo G. Corsico[1,4], James A. Irving[5], Annamaria Fra[2*], Ilaria Ferrarotti[1,4*]

[1]Centre for Diagnosis of Inherited Alpha-1 Antitrypsin Deficiency, UOC Pulmonology, Fondazione IRCCS Policlinico San Matteo, Pavia, Italy
[2]Experimental Oncology and Immunology, Department of Molecular and Translational Medicine, University of Brescia, Brescia, Italy
[3]Alpha-1 Foundation Ireland, Irish Centre for Genetic Lung Disease, Royal College of Surgeons in Ireland Education and Research Centre, Beaumont Hospital, Dublin, Ireland
[4] Department of Internal Medicine and Therapeutics, Pulmonology Unit, University of Pavia, Pavia, Italy
[5] UCL Respiratory, Rayne Institute and the Institute of Structural and Molecular Biology, University College London, London, UK

*Joint senior authors
Corresponding author:
Ilaria Ferrarotti Phone+39 0382 502620; Fax: +39 0382 502269
i.ferrarotti@smatteo.pv.it


Enhanced diagnostic strategy reveals novel variants


ORCID ID:
Stefania Ottaviani: 0000-0002-2039-230X
Giulia Bartoli: 0000-0001-8772-1852
Tomás P. Carroll: 0000-0002-0418-1641
Fabrizio Gangemi: 0000-0001-9389-651X
Alice M. Balderacchi 0000-0003-0714-1087
Valentina Barzon 0000-0003-1582-6887
Davide Piloni 0000-0002-7275-4229
Noel G. MacElvaney 0000-0002-0152-4370
Angelo G. Corsico 0000-0002-8716-4694

James A. Irving 0000-0003-3204-6356
Annamaria Fra: 0000-0002-4327-3004
Ilaria Ferrarotti: 0000-0003-4892-4192

**Keywords:** Serpins; serpinopathies; alpha-1-antitrypsin polymers; SERPINA1 rare variants; pathogenicity predictions.

**Abbreviations:** AAT, alpha-1-antitrypsin; AATD, alpha-1-antitrypsin deficiency; COPD, chronic obstructive pulmonary disease; ER, endoplasmic reticulum; PEI, polyethylenimine; SNP, single nucleotide polymorphism

**Web resources:** ClinVar: https://www.ncbi.nlm.nih.gov/clinvar/; gnomAD (Genome Aggregation Database): https://gnomad.broadinstitute.org/; dbsnp: https://www.ncbi.nlm.nih.gov/snp/; dbNSFP: http://database.liulab.science/dbNSFP; PolyPhen-2: http://genetics.bwh.harvard.edu/pph2/; PDBe: https://www.ebi.ac.uk/pdbe/.

This article has an online data supplement, which is accessible from this issue's table of content online at www.atsjournals.org.

**ABSTRACT**

Alpha-1-antitrypsin deficiency (AATD) is an under-diagnosed disorder associated with mutations in the *SERPINA1* gene encoding alpha-1-antitrypsin (AAT). Severe AATD can manifest as pulmonary emphysema and progressive liver disease. Besides the most common pathogenic variants S (E264V) and Z (E342K), many rarer genetic variants of AAT have been found in patients and in the general population. Here we report a panel of new *SERPINA1* variants including 4 null and 16 missense alleles identified among a cohort of individuals with suspected AATD whose phenotypic follow-up showed inconclusive or atypical results. As the pathogenic significance of the missense variants was unclear purely on the basis of clinical data, the integration of computational, biochemical and cellular studies was used to define the associated risk of disease. Established pathogenicity predictors and structural analysis identified a panel of candidate damaging mutations that were characterized by expression in mammalian cell models. Polymer formation, intracellular accumulation and secretory efficiency were evaluated experimentally. Our results identified two AAT mutants with a Z-like polymerogenic severe deficiency profile ($S_{milano}$ and $M_{campolongo}$) and three milder variants ($X_{sarezzo}$, $P_{dublin}$, $C_{tiberias}$). Overall, the experimentally determined behaviour of the variants was in agreement with the pathogenicity scores of the REVEL predictor, supporting the utility of this bioinformatic tool in the initial assessment of newly identified amino acid substitutions of AAT. Our study, in addition to describing 20 new SERPINA1 variants, provides a model for a multidisciplinary approach to classification of rare AAT variants and their clinical impact on individuals with rare AATD genotypes.

**INTRODUCTION**

Alpha-1-antitrypsin (AAT) is a circulating serpin (serine protease inhibitor) mainly synthesized and secreted by hepatocytes. The AAT protein is encoded by the *SERPINA1* gene within the protease inhibitor (PI) locus (14q31-32.3), which spans 12.2 kb and contains five exons and four introns (1). Pathogenic mutations in this gene cause the monogenic disorder alpha-1-antitrypsin deficiency (AATD) (OMIM #613490). People with AATD exhibit insufficient AAT activity in the plasma due to reduced circulating concentrations and/or reduced functional activity. The main role of AAT is inhibition of human neutrophil elastase (HNE), and its deficiencyleads to a substantially increased risk of lung diseases including emphysema, COPD and bronchiectasis, primarily due to the uncontrolled proteolytic activity of HNE on the lung parenchyma (2,3). Additionally, some AAT variants are present in the lung as inactive polymeric chains that either derive from circulating polymers secreted into the bloodstream by hepatocytes (4,5), or may form locally within the lungs as a result of local inflammation and exposure to cigarette smoke (6). AATD may also manifest with liver disease such as neonatal hepatitis, cirrhosis, and increased risk of hepatocellular carcinoma. Liver damage is mediated by conformational defects of the AAT mutants leading to formation of polymer chains (7), in the endoplasmic reticulum (ER) of hepatocytes and concomitant accumulation as inclusion bodies (8). Extrapulmonary manifestations of AATD include panniculitis and vasculitis, for which the underlying pathogenesis remains poorly understood (9,10).

An ever-increasing spectrum of mutations have been identified in the *SERPINA1* gene that yield AAT variants associated with deficiency and that exhibit dysfunctional behaviour. The canonical naming of AAT variants reflects their separation by isoelectric focusing (IEF), in which the first letters of the alphabet are used for anodal

migration and the last letters denote cathodal migration, occasionally appended by the birthplace of the index case (11,12). While convenient for phenotyping, the isoelectric point is not intrinsically reflective of pathogenicity. By convention, amino acid mutations are designated by the residue number in the mature secreted protein, lacking the 24 N-terminal residues of the signal peptide. This system however differs from that recommended by the Human Genome Variation Society (HGVS), in which numbering begins at the initiation methionine (13).

*SERPINA1* gene variants can be classified according to their plasma level and/or anti-elastase activity as normal, deficient, null or dysfunctional (1). A large number of these, including the most common M alleles (M1, M2, M3, M4, M5) and many rare variants, are characterized by AAT plasma levels that fall within general population reference ranges, and are not associated with an increased risk of lung or liver disease (14). Null variants, conventionally designated by 'Q0', are characterized by an absence of detectable AAT in the plasma by routine laboratory quantification and are associated with an increased risk of developing emphysema. In most cases these mutations are generated by premature stop codons, splicing site alterations, insertions or large deletions (15). An emerging class of AAT variants is found at the other extreme, comprised of a subset of coding variants whose plasma levels fall within the normal range but exhibit abnormal anti-protease function (16).

Of the pathogenic missense variants associated with AATD, S (E264V) and Z (E342K) are the most common, but there is also a wide spectrum of rare AAT variants that fall into this category (14,17). These deficiency variants are characterized by reduced AAT plasma levels, intrahepatic accumulation and are phenotypically associated with an increased risk of developing lung and/or liver disease. The pathogenetic mechanisms of intracellular accumulation have been mainly investigated for the Z AAT mutant. A

single mutation (E342K) in the breach region of the Z AAT molecule causes it to adopt a misfolded conformation in the ER of hepatocytes, making it a substrate of ER-associated degradation (ERAD) (18). A significant proportion of the misfolded Z AAT mutant escapes the quality-control degradation system and assembles through a domain-swap mechanism into long polymeric chains that condense as characteristic periodic acid–Schiff (PAS)-positive and diastase-resistant inclusion bodies within the ER (8). For the rare variants that have been investigated experimentally, such as Mmalton, Mprocida, Mwurzburg, Plowell, I, Pbrescia, the balance between secretion, degradation and aggregation determines their pathogenic potential.

Here we show the benefits of enhancement to a routine diagnostic pipeline with genotyping against an extended panel of disease-associated variants, selective gene sequencing, in silico pathogenicity prediction and selective cellular assays. We report 20 novel variants of the *SERPINA1* gene discovered through the Italian and Irish targeted detection programs for AATD. We applied established algorithms to predict those mutations most likely to affect the integrity of the AAT protein, and evaluated these variants in a mammalian cellular expression system. The distribution of the missense mutations across the protein highlights the susceptibility of the AAT structure to perturbation.Our study provides insights into the multidisciplinary approach required to achieve a precise diagnosis and classification of newly identified rare variants in AATD.

## MATERIALS AND METHODS

### Clinical data

Clinical data, obtained from direct observation of clinical charts and reported here in an anonymized form, are part of the Italian Registry of Alpha-1-antitrypsin Deficiency (RIDA1) (ethical approval number 0385 on 14-1-2019, by IRCCS Policlinico S.Matteo of Pavia) or the Irish National AATD Registry (Beaumont Hospital Research Ethics Committee number 05-03).

**Biochemical and molecular diagnosis of AATD**

Biochemical and genetic tests to diagnose AATD were utilized in this study with the understanding and written consent of each subject, according to each institution's ethical recommendations. Suspected AATD samples underwent a diagnostic algorithm comprising measurement of serum AAT and C-reactive protein (CRP), AAT phenotyping by IEF, genotyping for detection of S, Z and a panel of 14 rare AAT alleles (19) and, where inconclusive, *SERPINA1* gene sequencing by Sanger or NGS methods (1,20).

**Pathogenicity predictions and structural analysis**

The pathogenic impact of single amino acid variations resulting from missense SNPs was predicted computationally with PolyPhen-2 (21) and REVEL (14,22). Possible structural consequences were evaluated with reference to native AAT (a hybrid model of PDB 3NE4 and 1HP7).

**Sequence conservation analysis**

AAT orthologues 350-450 residues in length from a BLAST search (23) of the UniProtKB database (e-value ≤ 0.1) using entry P01009 were pruned with reference to a ClustalOmega (24) tree, with one complete sequence most similar to the human

protein retained for each species. Amino acid conservation at sites of interest within the resulting alignment was determined with the Python implementation of WebLogo (25).

**Expression vectors**

Expression vectors based on pcDNA3.1/Zeo (+) and encoding M1V and Z AAT were obtained previously (26). New vectors were obtained from M1V by the QuikChangeII site-directed mutagenesis kit. Reagents and primers are listed in Supplementary Tables S1 and S2.

**Cell transfection and lysis**

HEK 293T/17 (ATCC CRL-11268) and Hepa 1-6 (ATCC CRL-1830) cell lines were cultured in DMEM/10%FBS and transiently transfected by polyethylenimine (PEI) as described previously (27,28) and in the Online Data Supplement. Transfected cells were incubated for 20 h in Opti-MEM. Culture media were then collected and centrifuged to remove cell debris, while the cells were lysed in a 1% NP40 buffer. NP40-soluble fractions were separated by centrifugation from the insoluble fractions, which were resuspended in equal volumes of lysis buffer and solubilized by sonication.

**SDS-PAGE and immunoblots**

Samples in SDS-sample buffer/25mM DTT were separated by 8% SDS-PAGE and transferred onto LF-PVDF membranes, subsequently probed with anti-human AAT polyclonal antibody (Dako) and HRP-conjugated anti-rabbit IgG antibody revealed by ECL Plus.

## Quantification of AAT

Total and polymeric AAT were quantified by two sandwich ELISA protocols, as described in the Online Data Supplement.

## Immunofluorescence and confocal microscopy

Hepa 1-6 cells grown on glass coverslips overnight were transfected, fixed with 4% paraformaldehyde after 24h, permeabilized (5 min in PBS/0.1% TritonX100), saturated with PBS/0.1% BSA/0.05% TritonX100 and immunostained with anti-AAT polymer 2C1 mAb (0.5µg/mL) (29), followed by rabbit anti-human AAT (Dako) (2 µg/mL), and AlexaFluor-conjugated secondary antibodies. Nuclei were stained with DAPI. Coverslips were mounted with ProLong and analysed on a Zeiss LSM510 confocal microscope.

## RESULTS

## Novel SERPINA1 allelic variants

During the diagnosis of individuals referred to the Centre for Diagnosis of Inherited Alpha-1-Antitrypsin Deficiency (Pavia, Italy) or the Alpha-1 Foundation Ireland, Beaumont Hospital (Dublin, Ireland) for suspected AATD, we identified 20 novel allelic variations in the SERPINA1 gene, comprising 4 null mutations and 16 missense single nucleotide polymorphisms (SNPs) (Table1). None had been previously described. Genetic, biochemical and clinical data on the probands are reported in the Online Data Supplement and Table S3 for each variant and are summarized in Table 1.

**Pathogenicity predictions and structural analysis**

The occurrence of newly identified AAT variants in simple or compound heterozygosity and their variable presentation substantially limits the interpretation of their phenotypic impact on the carriers. Several bioinformatic tools have been developed to help predict the functional effects of coding variations, based on analysis of conservation in multiple sequence alignments, the biochemical properties of amino acid substitutions and the impact on 3D structures of the affected protein.

We subjected the 16 newly identified AAT missense variations to two pathogenicity prediction algorithms, Polyphen-2 (21) and REVEL (22), to predict their pathogenic significance and to prioritize them for characterization in cellular models (Table 2). While Polyphen-2 is widely used, the REVEL algorithm emerged as the best-performing predictor of pathogenicity for *SERPINA1* non-synonymous polymorphisms (14); in that work, Giacopuzzi and co-authors defined three groups that optimize classification based on cluster analysis of REVEL scores: cluster 1 identifies the likely-benign AAT variants (0–0.354), cluster 3 the probably pathogenic ones (0.618–1), while cluster 2 (0.355–0.618) includes variants with milder or uncertain significance. Applying these cut-off values, the majority of the new variants were predicted as neutral (score <0.355), two variants (E98K and K290Q) produced intermediate scores (0.355 to 0.618), and three AAT variants (Y244H, D341Y and L383P) were classified as deleterious (score >0.618). Polyphen-2, which classifies as benign the substitutions with a score below 0.15, provided similar results, but also predicted the E205K variant as possibly damaging.

To extend our analysis, the localization and intra-molecular interactions of the mutated residues were evaluated using the native AAT structure (Figure 1A) with additional

reference to patterns of conservation among putative AAT orthologues across 82 species (Table 2).

*Truncation variants*

Three out of four identified null variants are the result of mutations that cause a loss of key C-terminal structural elements. P326Ter (Q0$_{sansiro}$) lacks strands 5A, 1C, 4B, 5B and the reactive centre loop. In a study of the folding properties of AAT fragments, one that incorporated a near-identical residue range (1-323) was found to be soluble and behave as a molten globule folding intermediate (30), and therefore in a cellular context it may elicit the unfolded protein response and would certainly contribute to the degradative burden. Both K368Ter (Q0$_{firenze}$) and P369Pfs*5 (Q0$_{foggia}$) lack strands 4B and 5B that are required for positioning of the reactive centre loop and strand 1C in the native state. There is evidence that an AAT molecule lacking an incorporated C-terminus is unable to undergo the self-insertion process (30) that is believed to be central to the polymerization mechanism (8). By analogy with the heteropolymerization of Z with M and S variants, it is conceivable that a proportion of these molecules may capture co-expressed variants as heterodimers (31,32).

Due to the intronic localization of the Q0$_{siracusa}$ (c.-5+1800_1804delC), we cannot provide a similar analysis. Since this nevertheless results in an absence of AAT, we conclude that the deletion may directly affect a splicing site.

*Likely benign variants*

The variants predicted to be benign are largely surface-exposed, mediating either no or minimal observed interactions with other structural elements, and in sites generally found to be polymorphic in orthologues across species (Supplementary Figure S1). *Q393K (X$_{magenta}$)* and T22A (M$_{asti}$) are located in regions of the AAT structure that are

disordered in at least some crystal structures and are at sites tolerant of substitutions in orthologues (Table 2). Similarly, H43Q (M$_{napoli}$), T150A (M$_{monza}$), D270N (L$_{bressanone}$), T273N (M$_{ancona}$), S301R (V$_{verceia}$) are at positions that do not make important contacts with other residues. Others show conservative polar substitutions that are expected to retain similar interactions. Position 133, sited within helix E, forms a polar contact with K129 which would be expected to be maintained in the D133N (P$_{savona}$) variant; T268N (M$_{andria}$) is similarly predicted to maintain a network of polar interactions at the N-terminal end of helix H. The substitution in this category most likely to have a structural impact, K222T (F$_{milano}$), results in a loss of a salt bridge that contributes to the interface between 3C and 4C, but the site is generally tolerant of polar substitutions among orthologues and is thus unlikely to significantly perturb AAT structure.

*Potentially damaging variants*

Four of the six variants with high REVEL/Polyphen scores all arise in the vicinity of the AAT 'breach' region (Figure 1B). K290Q (C$_{tiberias}$) is proximate to the location of the Z mutation, E342; whilst it had been hypothesized in early studies that loss of a 290-342 salt bridge would contribute to the polymerogenic behaviour of Z, the partial tolerance of this site to mutagenesis is demonstrated by a K290E mutation that reduced secretion in a cell model by only around a third (33). This modest effect is consistent with its intermediate REVEL score.

The strictly conserved Y244 side chain is integral to the 'breach' region at the top of β-sheet A that is the site of incorporation of the reactive centre loop during the inhibition of a target protease. Destabilization of this region in the presence of the pathogenic Z mutation causes AAT to adopt the M* intermediate with solvation of a cryptic pocket (34). In the Y244H (P$_{dublin}$) variant, the histidine side-chain would

maintain a polar character but perturb the water-bonding network of the breach (Figure 1B). A molecular dynamics simulation, performed to investigate the possible consequences in more detail, suggests that changes to the local interaction network, including new interactions that are established between the NH group on the ring of H244 and the backbone atoms of G192 and K193, likely are the cause of a structural alteration of a strand of β-sheet A (Supplementary Figures S2 and S3).

*D341Y (S$_{milano}$)* also represents substitution of a strictly conserved residue. Loss of this residue disrupts a network of water-mediated interactions at a site directly preceding position 342 of the Z mutation (Figure 1B). This might be expected to partially destabilise the top of strand 5A, a factor which is known to result in polymerisation (34). Indeed, in molecular dynamics simulations, analysis of the network of local interactions of the wild-type protein shows that D341, as well as E342, due to its negatively charged side chain, has a crucial role in the stability of the upper end of strands s5A and s6A in β-sheet A as well as the hinge region between residues 340 and 344 (Supplementary Figure S4).

L383P (M$_{campolongo}$), situated in β-strand 5B, is buried adjacent to the 'breach' region and the cryptic pocket that is revealed in the M* intermediate (35,36). The replacement of leucine by proline would alter the hydrophobic packing of this region, and the stereochemical constraints imposed on the backbone by the proline residue would additionally be expected to partially destabilize the local β-strand conformation. As this element is involved in an intermolecular domain swap during C-terminal-mediated polymerization (7), a destabilized C-terminus could result in displacement of this region and therefore polymer formation.

Inspection of the E98K (P$_{darfo}$) and E205K (X$_{sarezzo}$) variants makes their relatively high scores somewhat surprising, as in the native structure they are surface-exposed and

lack a clear interaction with other residues, although they would alter the local surface charge.

These observations collectively suggest that the mutations most likely to result in pathogenic outcomes are situated in the vicinity of the breach region. We chose these four residues, as well as the two others with intermediate pathogenicity scores, for further experimental characterization.

## Characterization of the AAT variants expressed in cellular models

*Intracellular polymerization*

The variants selected by our preliminary bioinformatic and structural evaluation were expressed and characterized in both the HEK293T and the Hepa 1-6 cell lines, with wild-type M and the polymerogenic Z AAT mutants used for reference. Polymer formation was first assessed by the tendency of the proteins to accumulate as intracellular aggregates resistant to NP40 treatment (37).

Cells were transiently transfected and cell media and extracts that were soluble or insoluble by 1% v/v NP40 treatment, were evaluated by SDS-PAGE and immunoblot (Figure 2). As expected for a readily secreted protein, M AAT was exclusively present in the detergent-soluble fraction, and a similar profile was observed for the E98K AAT variant in both cell lines. Polymerogenic Z AAT was distributed approximately equally between the NP40-soluble and -insoluble fractions and a similar profile was observed for the D341Y and L383P variants. Traces of NP40-insoluble polymers were detected in the E205K, Y244H and K290Q variants. Analysis of the AAT variants in the cell

media showed marked deficiency of Z as well as of the polymerogenic D341Y and L383P.

*Secretion profile*

Impact of the mutations on secretion efficiency was evaluated by quantifying AAT in the culture media of both HEK293T and Hepa1.6 transfected cells by sandwich ELISA (Figure 3). Consistent with previous studies (26,31,38), the Z variant was secreted significantly less than wild-type M AAT. In comparison, E98K was secreted at nearly normal levels, L383P showed the most severe reduction, a moderate reduction was observed for D341Y, while a milder but significant deficiency was observed for the other variants. Strikingly, when the secretion levels of the AAT variants were plotted against their REVEL scores (Figure 3, right panels), a common trend was observed supporting the potential of this algorithm to prospectively predict deficiency and associated disease risk.

**Accumulation of polymerogenic mutants within inclusion bodies**

A fraction of polymerogenic variants can be secreted as polymers (5). To determine whether there was any direct evidence of polymer formation by the novel variants we evaluated their presence in the cell media by a sandwich ELISA based on the monoclonal antibody 2C1. This experiment revealed detectable polymer formation by D341Y and L383P (Figure 4A).

Intracellular accumulation and distribution of these polymerogenic AAT variants was subsequently investigated by immunofluorescence staining and confocal microscopy.

Wild-type M, the polymerogenic Z AAT and D341Y and L383P were transiently expressed in Hepa1.6 cells and the cells were fixed and stained with the polymer-specific 2C1 mAb and the anti-total AAT polyclonal antibody (Figure 4B). As expected, wild-type M AAT was only recognized by the non-conformation selective polyclonal antibody, with a reticular staining pattern and a perinuclear localization, consistent with accumulation within the Golgi apparatus. Z AAT resulted in the formation of 2C1-positive punctate structures resembling the inclusion bodies previously reported in cellular models expressing Z AAT (29,39). Similar patterns were seen for the D341Y and L383P variants, thus confirming their propensity to accumulate Z-like, 2C1-positive polymers within the cell.

**DISCUSSION**

With developments in clinical pipelines and advances in genomic technologies, an ever-increasing number of rare AAT variants are being identified. Many rare variants are not easily diagnosed with routine laboratory techniques, thus contributing to misclassification and misdiagnosis. For example, phenotyping by IEF usually does not allow the precise identification of M-like AAT variants, such as $M_{malton}$, $M_{procida}$ (40) as they are easily confused with M (normal) protein. Furthermore, Null (Q0) variants in heterozygosity are invisible to IEF and can only be indirectly observed through their effects on circulating levels. Most rare alleles can only be detected by molecular biology techniques, such as *SERPINA1* gene sequencing (1, 41). Sometimes, depending on the complexity of the case, sequencing of exons might not be sufficiently informative and more extensive analysis by whole gene sequencing is required to detect rare exonic deletions, alterations in promoters or deep intronic regions not covered by standard sequencing (9,40). It has been recently demonstrated that the

choice of diagnostic algorithm can have a significant impact on the accurate diagnosis of AATD, which is essential for appropriate medical care and treatment (42). This is exemplified by the finding of the $Q0_{foggia}$ mutation, described herein, that is *in cis* with an S allele: a simpler diagnostic algorithm, that includes only the detection of most common variants, would have identified a low-risk S variant rather than a high-risk Null allele.

Precise identification of the AAT genotype facilitates the accurate clinical management of a patient, but for variants that are uncharacterized or with conflicting evidence for pathogenicity this must be interpreted through the lens of molecular pathological mechanisms: diagnostic procedures should provide this combined information to clinicians. Virtually all *SERPINA1* nonsense mutations will have a deleterious effect due to lack of protein synthesis, severe effects on the folding and stability of the protein, or lack of activity: there is no dispensable C-terminal component to the AAT structure. Therefore, individuals with a novel Null variant should always be considered at particularly high risk of emphysema within the spectrum of AATD, with a similar prognosis to other previously reported Null alleles (15,43).

Unlike Null, the molecular mechanisms of missense variants are more variable and the pathogenicity of novel missense variants cannot be inferred by clinical data gathered from a single individual (44,45). The presumption of a 'benign' or 'pathogenic' mutation without robust supporting evidence, risks inappropriate management of AATD. Mis-interpretation of novel variant(s) can lead to inappropriate diagnostic calls and clinical care.

Several bioinformatic tools have been developed to streamline the functional assessment of newly identified protein coding variants, based on multiple features such as amino acid or nucleotide conservation, the biochemical properties of amino

acid substitutions and structural data. REVEL is an ensemble predictive method which performed best in discriminating known pathogenic AAT variants from those classified as benign (14). In the present study we found that REVEL showed good correlation with the level of secretion exhibited by variants in the two mammalian expression systems used.

Cellular models of disease have proven to be good correlates with clinical manifestations of deficiency alleles, by recapitulating the consequences of misfolding, degradation and aggregation in a native-like environment. This is generally accomplished by transient transfection in hepatoma cells or in non-hepatic cell lines achieving higher expression levels, such as HEK293T or COS-7 cells (27,28,36,46). Here, we used both Hepa1.6 and HEK293T cell lines which revealed similar behaviour for each of the variants evaluated: Y244H (P$_{dublin}$), D341Y (S$_{milano}$) and L383P (M$_{campolongo}$) showed moderate to severe secretion deficiency, with the latter two also exhibiting marked accumulation of intracellular NP40-insoluble polymers and 2C1-positive inclusions by confocal microscopy. A milder deficiency was observed for E205K (X$_{sarezzo}$) and K290Q (C$_{tiberias}$). Our results suggest that individuals with D341Y and L383P variants in heterozygosity with other deficiency variants may have an associated risk of developing both lung and liver dysfunction. The detailed characterization presented here further extends the number of naturally occurring rare mutations in AAT with pathogenic potential that have been identified.

Naturally arising serpin variants have played a key role in the elucidation of mechanisms underlying molecular dysfunction including polymerization (47). Of the newly described AAT variants, those that are not expected to exhibit a pathogenic

phenotype are for the most part surface-exposed, polar, not involved in interactions with other structural elements, and at sites that are relatively polymorphic among orthologues in other species. Our structural analyses reveal that collectively, the variants occupy a diversity of positions throughout the molecule; it is therefore striking that those we show to be the most deleterious all occur in proximity to the 'breach'. This region of the molecule is destabilized in the Z variant (7) and overlies the C-terminal strands s4B and s5B that are displaced to form intermolecular interactions during polymer formation in the liver (8). In addition, the breach is the point of insertion of the reactive centre loop during inhibition of a target protease, and, by analogy with the Z variant, these mutations may also result in a reduction of protease inhibitory efficiency.

The results presented here show that routine incorporation of genetic analysis within an AATD diagnostic pipeline can provide important clinical insight that would otherwise escape detection. In the event of an uncharacterized variant being identified, these data further support the usefulness, in the absence of other experimental information, of pathogenicity predictors in establishing the likelihood of a deleterious effect on the protein. This can be augmented by a consideration of the structural context in relation to deficiency 'hot spots' in the vicinity of the breach and shutter domains. Ultimately, where possible, the gold-standard approach in a research setting remains evaluation of the variant in a cellular context.

Other members of the serpin superfamily are associated with conformational diseases, called serpinopathies, associated with misfolding and ordered aggregation, causing conditions ranging from dementia to thrombosis. Our results indicate that

pathogenicity predictors are compatible with the idiosyncrasies of the serpin fold and accordingly our approach - combining an in-silico triage with subsequent evaluation of cellular accumulation and export – has the potential to benefit diseases arising from molecular defects in related proteins. In a wider context, the paper provides evidence of the usefulness of integration of cellular models, prediction analysis and clinical data to guide the care of patients with genetic diseases caused by uncharacterized rare mutations.

**ACKNOWLEDGEMENTS**

| Variant name | Nucleotide change (HGVS) | Amino acid change (HGVS) | Amino acid change (conventional) | Base allele | dbSNP |
|---|---|---|---|---|---|
| Q0$_{siracusa}$ | c.-5+1800_1804delC | | | M1(Val) | |
| Q0$_{sansiro}$ | c.1621_1624delC | p.Pro350Ter | P326Ter | M1(Val) | |
| Q0$_{firenze}$ | c.1747A>T | p.Lys392Ter | K368Ter | M1(Val) | |
| Q0$_{foggia}$ | c.1750_1752delC | p.Pro393Pro | P369Pfs*5 | S | |
| M$_{asti}$ | c.709A>G | p.Thr46Ala | T22A | M2 or M2obernburg | rs1243707738 |
| M$_{napoli}$ | c.774C>G | p.His67Gln | H43Q | M1(Val) or M3 | rs140744031 |
| P$_{darfo}$ | c.937G>A | p.Glu122Lys | E98K | M1(Val) or M1(Ala) | |
| P$_{savona}$ | c.1042G>A | p.Asp157Asn | D133N | M1(Val) | |
| M$_{monza}$ | c.1093A>G | p.Thr174Ala | T150A | M1(Val) | rs1450215301 |
| X$_{sarezzo}$ | c.1258G>A | p.Glu229Lys | E205K | M1(Val) | |
| F$_{milano}$ | c.1310A>C | p.Lys246Thr | K222T | M1(Val) | rs864622054 |
| P$_{dublin}$ | c.1375T>C | p.Tyr268His | Y244H | M1(Val) or M3 | |
| M$_{andria}$ | c.1448C>A | p.Thr292Asn | T268N | M1(Val) | |
| L$_{bressanone}$ | c.1453G>A | p.Asp294Asn | D270N | M1(Val) | rs772436715 |
| M$_{ancona}$ | c.1463C>A | p.Thr297Asn | T273N | M1(Val) | rs752776707 |
| C$_{tiberias}$ | c.1513A>C | p.Lys314Gln | K290Q | M1(Val) | |
| V$_{verceia}$ | c.1548C>G | p.Ser325Arg | S301R | M1(Val) | |
| S$_{milano}$ | c.1666G>T | p.Asp365Tyr | D341Y | M3 | rs143370956 |
| M$_{campolongo}$ | c.1793T>C | p.Leu407Pro | L383P | M1(Ala) or M2 | |
| X$_{magenta}$ | c.1822C>A | p.Gln417Lys | Q393K | M2 | |

Table 1- List of novel variants described in the present paper. The reference sequence used for the *SERPINA1* transcript annotation is NCBI NM_001127701.1, while the reference sequence for the SERPINA1 protein is Uniprot P01009. Variants are sorted by gene location.

| Amino acid change (HGVS) | REVEL | PolyPhen-2 Hvar-HDiv | Expo-sure | Domain | Conser-vation | Structural features |
|---|---|---|---|---|---|---|
| Null | | | | | | |
| P326Ter | n/a | n/a | - | Δs5A/ RCL/ 1C/ 4B/ 5B/ | | Likely to form a soluble, non-native molten globule conformation** |
| K368Ter | n/a | n/a | - | Δs4B/ 5B | | Missing the region exchanged in a C-terminal polymer; feasible a proportion may form heterodimer by accepting C-terminus from a full-length variant |
| P369Pfs*5 | n/a | n/a | - | Δs4B/ 5B | | As above |
| Predicted benign | | | | | | |
| T22A | 0.221 | 0/ 0 | E | N-terminus | | In a region disordered in most 3D structures; alanine predominant in other species |
| H43Q | 0.264 | 0.002-0.005 | E | hA | | Mutation-tolerant site |
| D133N | 0.276 | 0.004-0.033 | E | hE | | Conservative change expected to maintain weaker polar contact with K129; change in surface charge |
| T150A | 0.285 | 0.001-0.005 | E | hF | | Sidechain does not make important contacts; mutation-tolerant site; alanine tolerated in other species |
| K222T | 0.196 | 0.001-0.162 | E | s3C | | Loss of stabilising salt bridge between strands 3C/4C; site tolerant of polar substitutions in other species |
| T268N | 0.240 | 0.018-0.216 | E | | | Polar substitution at a site that stabilizes helix H by a hydrogen bond; asparagine seen in other species so conservative change likely maintains contacts |
| D270N | 0.167 | 0.080-0.083 | E | hH | | Conservative substitution; sidechain does not make important contacts |
| T273N | 0.085 | 0,003-0.109 | E | hH | | Sidechain does not make important contacts |
| S301R | 0.227 | 0-0 | E | hI | | Sidechain not involved in important contacts; change in surface charge |
| Q393K | 0.161 | 0.004-0.033 | E | | | Polar substitution of sidechain that does not make important contacts; change in surface charge |
| Predicted damaging | | | | | | |
| E98K | 0.392 | 0.004-0.015 | E | hD | | Breaks intra-helix D salt bridge, minimal effects on stability expected; polar amino acid mutation-tolerant site; change in surface charge |
| E205K | 0.217 | 0.219-0.169 | E | s4C | | Sidechain does not make important contacts; change in surface charge |

| | | | | | | |
|---|---|---|---|---|---|---|
| Y244H | 0.798 | 1-1 | B | s2B | | Packs against conserved positions 190 and 194; charged polar replacement expected to destabilize the functionally important breach region: tendency towards polymerization and some effects on inhibitory activity might be expected; highly conserved site. |
| K290Q | 0.569 | 0.326-0.948 | E | | | Highly conserved site; loss of salt bridge to conserved position 342; K290E mutation at this site did reduce secretion** with effect on polymerization unknown |
| D341Y | 0.648 | 0.002-0.005 | E | s5A | | Located immediately C-terminal to strand 5A in a region sensitive to destabilization, associated with polymer formation**; loss of highly conserved residue |
| L383P | 0.819 | 0.004-0.015 | B | s5B | | Proline forms β-bulge in strand 5B, destabilization of which may promote C-terminal polymerization |

- Table 2. Novel AAT variants and predicted impacts. ** See refs. 30, 33 and 34

| Variant name | Mutation | RefSNP | patient code | Genotype (PI*) | AAT level (mg/dL) | CRP level (mg/dL)* | Age (years) | P/y | Clinical manifestations | Index cases and relatives |
|---|---|---|---|---|---|---|---|---|---|---|
| Q0siracusa | c.-5+1800_1804delC | | A1 | M1/Q0siracusa | 88 | 0,19 | 71 | 0 | dyspnea | index case |
| | | | A2 | M1/Q0siracusa | 81,3 | 0,05 | 57 | 20,5 | emphysema/COPD | |
| | | | A3 | M1/Q0siracusa | 92,2 | 0,09 | 57 | 45 | emphysema | |
| | | | A3.1 | M1/Q0siracusa | 125 | 0,11 | 59 | 40 | healthy | brother of A3 |
| | | | A3.2 | M1/Q0siracusa | 113,4 | 0,29 | 48 | 0 | healthy | brother of A3 |
| | | | A3.3 | M1/Q0siracusa | 155,1 | 0,27 | 55 | 25 | healthy | brother of A3 |
| | | | A3.4 | M1/Q0siracusa | 177,1 | 0,07 | 19 | 0 | healthy | daughter of A3 |
| | | | A3.5 | M1/Q0siracusa | 139,4 | 0,07 | 22 | un | healthy | son of A3 |
| Q0sansiro | P326Ter | | B1 | Mwurzburg/Q0sansiro | 38,5 | 0,09 | 33 | 0 | healthy | index case |
| Q0firenze | K368Ter | | C1 | M1/Q0firenze | 78,7 | 0,07 | 49 | un | healthy | index case |
| | | | C1.1 | M3/Q0firenze | 76 | 0,15 | 23 | 0,25 | healthy | son of C1 |
| Q0foggia | P369Pfs*5 | | D1 | M1/Q0foggia | 81 | un | 46 | 0 | chronic bronchitis/asthma | index case |
| | | | D1.1 | M1/Q0foggia | 69 | 0,43 | 76 | un | un | father of D1 |
| Masti | T22A | rs1243707738 | E1 | M2obernburg/Masti | 102 | 0,22 | 57 | 0 | bronchitis/bronchiectasis | index case |
| Mnapoli | H43Q | rs140744031 | F1 | M1/Mnapoli | 91,2 | 0,11 | 79 | ex | bronchitis/emphysema/dyspnea | index case |
| | | | F2 | M1/Mnapoli | 113,9 | 0,03 | 56 | 40 | emphysema | |
| Pdarfo | E98K | | G1 | M1/Pdarfo | 250 | 2,57 | 83 | un | vascular pathology | index case |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Psavona | D133N | | H1 | M2/Psavona | 172 | un | 46 | un | haemophilia | index case |
| Mmonza | T150A | rs1450215301 | I1 | M2/Mmonza | 117,7 | 0,08 | 45 | 42 | emphysema/dyspnea | index case |
| Xsarezzo | E205K | | J1 | M3/Xsarezzo | 105 | un | 42 | un | healthy | index case |
| | | | J2 | M1/Xsarezzo | 106 | un | 64 | un | healthy | |
| | | | J3 | M2/Xsarezzo | 137,6 | 0,06 | 14 | un | un | |
| | | | J4 | M1/Xsarezzo | 173,3 | 0,06 | 69 | 0 | asthma | |
| | | | J4.1 | M1/Xsarezzo | 168,4 | 0,07 | 40 | 0 | healthy | daughter of J4 |
| Fmilano | K222T | rs864622054 | K1 | M1/Fmilano | 194,1 | 0,05 | 48 | 0 | asthma/emphysema | index case |
| | | | K2 | M1/Fmilano | 126,9 | 0,14 | 53 | un | emphysema | |
| | | | K3 | M1/Fmilano | 78,3 | 0,03 | 50 | 40 | un | |
| Lmestre | S236F | | L1 | M1/Lmestre | 224 | 1,22 | 52 | un | un | index case |
| Pdublin | Y244H | | M1 | M3/Pdublin | 181 | un | 60 | un | cough/wheeze | index case |
| Mandria | T268N | | N1 | Xchristchurch/Mandria | 88 | 0,6 | 50 | 6,7 | asthma/bronchitis/emphysema | index case |
| Lbressanone | D270N | rs772436715 | O1 | Z/Lbressanone | 96,7 | 2,71 | 18 | un | un | index case |
| | | | O1.1 | M1/Lbressanone | 131 | 0,04 | 51 | 0 | healthy | mother of O1 |
| Mancona | T273N | rs752776707 | P1 | Z/Mancona | 66 | 0,05 | 63 | un | un | index case |
| Ctiberias | K290Q | | Q1 | M1/Ctiberias | 137 | un | 53 | un | abnormal liver enzymes/dyspnea | index case |
| | | | Q2 | M2/Ctiberias | 105 | 0,41 | 80 | un | bronchiectasis | |
| Vverceia | S301R | | R1 | M1/Vverceia | 132,6 | 0,06 | 50 | 0 | healthy | index case |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Smilano | D341Y | rs143370956 | S1 | M1/Smilano | 119 | 0,02 | 44 | 0 | bronchiectasis | index case |
| Mcampolongo | L383P | | T1 | M2/Mcampolongo | 72 | 0,99 | 58 | un | un | index case |
| Xmagenta | Q393K | | U1 | M1/Xmagenta | 121 | 0,21 | 56 | un | un | index case |
| | | | U1.1 | M1/Xmagenta | 123 | 0,42 | 30 | un | un | son of U1 |
| | | | U1.2 | M1/Xmagenta | 157 | un | 28 | un | un | son of U1 |

*CRP normal value <0.8 mg/dL

(un=unknown; p/y= pack/year)

Table 3

**FIGURE LEGENDS**

**Figure 1. Location of variants identified in this study with reference to native AAT.** (A) Predicted minimally-perturbing (red) and perturbing (black) variants, mapped onto the structure of AAT reconstructed from PDB entries 1HP7 and 3NE4. (B) Detailed views of residues predicted to be perturbing are also shown, in which the wild-type residue is shown in dark yellow stick, the mutation in magenta stick, nearby residues are pale yellow, and water molecules red spheres. Polar interactions present in both the wild-type and mutant are shown as dashed blue lines, those present only in the wild-type are red, and those only in the mutant are yellow; black dashes indicate an altered but preserved bond. Secondary structure elements are β-sheets A (blue), B (green) and C (yellow), with helices in cyan and loop regions grey.

**Figure 2.** Characterization of AAT variants in cell models. HEK293T (A) and Hepa 1-6 (B) cells were transfected to express the indicated AAT variants and incubated for 24h in serum-free Opti-MEM. The cell culture media were collected, and the cells lysed in 1% v/v NP40 buffer. NP40-soluble material was separated by centrifugation from insoluble material containing nuclei and AAT polymer aggregates. The latter was resuspended and sonicated, and both cellular fractions were separated in 8% SDS-PAGE and detected by immunoblot with a polyclonal anti-AAT antibody (Dako). Black and white arrowheads indicate high mannose and complex N-glycosylated forms of AAT, respectively.

**Figure 3.** Secretion efficiency of the novel AAT variants. Total AAT levels in the cell media of HEK293T (A) or Hepa 1-6 (B) transfected cells were measured by ELISA and represented as percentage of the wild-type M levels (left panels) (mean±SD, with experimental values as blue circles; n=2 and n=3 independent transfection experiments for HEK293T for Hepa 1-6, respectively) (1-way ANOVA, $p<0.0001$; Dunnett's multiple comparison test between each variant and M AAT, * $p<0.05$, ** $p< 0.01$, *** $p< 0.001$, **** $p<0.0001$). Secretion AAT levels for each variant was plotted versus the REVEL score assigned to the variant (right panel) and correlation was determined by linear regression ($R^2=0.53$ for HEK293T and $R^2=0.57$ for Hepa 1-6).

**Figure 4.**

Mutant variants of AAT accumulate 2C1-positive polymers. (A) Polymeric AAT in the cell media of transfected HEK293T (top) and Hepa 1-6 (bottom) cells were measured by a sandwich ELISA based on mAb 2C1 capture, normalized to total AAT levels and represented as percentage of Z AAT polymer levels (mean±SD with experimental values as blue circles, n=2 independent transfections for both HEK293T and Hepa 1-6). (B) Hepa 1-6 cells seeded on glass coverslips were fixed with 4% paraformaldehyde 24 h after transfection with the wild-type M AAT, the polymerogenic mutant Z AAT, D341Y and L383P AAT variants. After permeabilization, cells were stained with the anti-AAT polymer 2C1 mAb (red) and a rabbit anti-human AAT pAb (Dako) (green), detected with goat anti-mouse IgG secondary antibody conjugated to Alexa Fluor 594 and goat anti-rabbit antibody conjugated to Alexa Fluor 488, respectively. Images were acquired on a Zeiss LSM510 confocal microscope with a 63X objective (1.4 oil). Scale bar: 20 µm.

# REFERENCES

1.  Ferrarotti I, Ottaviani S. Laboratory diagnosis. In: Strnad P, Brantly ML, Bals R, editors. α1-Antitrypsin deficiency (ERS monograph). Sheffield: European Respiratory Society; 2019. p. 39-51.

2.  Greene CM, Marciniak SJ, Teckman J, Ferrarotti I, Brantly ML, Lomas DA, Stoller JK, McElvaney NG. α1-Antitrypsin deficiency. Nat Rev Dis Primers 2016;2:16051.

3.  Strnad P, McElvaney NG, Lomas DA. Alpha1-Antitrypsin Deficiency. N Engl J Med 2020;382:1443-1455.

4.  Tan L, Dickens JA, Demeo DL, Miranda E, Perez J, Rashid ST, Day J, Ordoñez A, Marciniak SJ, Haq I, Barker AF, Campbell EJ, Eden E, McElvaney NG, Rennard SI, Sandhaus RA, Stocks JM, Stoller JK, Strange C, Turino G, Rouhani FN, Brantly M, Lomas DA. Circulating polymers in α1-antitrypsin deficiency. Eur Respir J 2014;43:1501-4.

5.  Fra A, Cosmi F, Ordoñez A, Berardelli R, Perez J, Guadagno NA, Corda L, Marciniak SJ, Lomas DA, Miranda E. Polymers of Z α1-antitrypsin are secreted in cell models of disease. Eur Respir J 2016;47:1005-9.

6.  Alam S, Li Z, Janciauskiene S, Mahadeva R. Oxidation of Z α1-antitrypsin by cigarette smoke induces polymerization: a novel mechanism of early-onset emphysema. Am J Respir Cell Mol Biol. 2011 ;45:261-9.

7.  Yamasaki M, Sendall TJ, Pearce MC, Whisstock JC, Huntington JA. Molecular basis of α1-antitrypsin deficiency revealed by the structure of a domain-swapped trimer. EMBO Rep. 2011;12(:1011-7.

8.  Faull SV, Elliston ELK, Gooptu B, Jagger AM, Aldobiyan I, Redzej A, Badaoui M, Heyer-Chauhan N, Rashid ST, Reynolds GM, Adams DH, Miranda E, Orlova EV, Irving JA, Lomas DA. The structural basis for Z $\alpha_1$-antitrypsin polymerization in the liver. Sci Adv. 2020;6:eabc1370.

9. Franciosi AN, Ralph J, O'Farrell NJ, Buckley C, Gulmann C, O'Kane M, Carroll TP, McElvaney NG. Alpha-1 antitrypsin deficiency-associated panniculitis. J Am Acad Dermatol 2021:S0190-9622.

10. Lyons PA, Rayner TF, Trivedi S, Holle JU, Watts RA, Jayne DR, Baslund B, Brenchley P, Bruchfeld A, Chaudhry AN, Cohen Tervaert JW, Deloukas P, Feighery C, Gross WL, Guillevin L, Gunnarsson I, Harper L, Hrušková Z, Little MA, Martorana D, Neumann T, Ohlsson S, Padmanabhan S, Pusey CD, Salama AD, Sanders JS, Savage CO, Segelmark M, Stegeman CA, Tesař V, Vaglio A, Wieczorek S, Wilde B, Zwerina J, Rees AJ, Clayton DG, Smith KG. Genetically distinct subsets within ANCA-associated vasculitis. N Engl J Med 2012;367:214-23.

11. Fagerhol MK, Laurell CB. The Pi system-inherited variants of serum alpha 1-antitrypsin. Prog Med Genet 1970;7:96-111.

12. Cox DW, Johnson AM, Fagerhol MK. Report of Nomenclature Meeting for alpha 1-antitrypsin, INSERM, Rouen/Bois-Guillaume-1978. Hum Genet. 1980;53:429-33.

13. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M, Lyon E, Spector E, Voelkerding K, Rehm HL; ACMG Laboratory Quality Assurance Committee. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med 2015;17:405-24.

14. Giacopuzzi E, Laffranchi M, Berardelli R, Ravasio V, Ferrarotti I, Gooptu B, Borsani G, Fra A. Real-world clinical applicability of pathogenicity predictors assessed on SERPINA1 mutations in alpha-1-antitrypsin deficiency. Hum Mutat. 2018;39:1203-1213.

15. Ferrarotti I, Carroll TP, Ottaviani S, Fra AM, O'Brien G, Molloy K, Corda L, Medicina D, Curran DR, McElvaney NG, Luisetti M. Identification and characterisation of eight novel SERPINA1 Null mutations. Orphanet J Rare Dis. 2014;9:172.

16. Laffranchi M, Elliston ELK, Gangemi F, Berardelli R, Lomas DA, Irving JA, Fra A. Characterisation of a type II functionally-deficient variant of alpha-1-antitrypsin discovered in the general population. PLoS One. 2019;14:e0206955.

17. Seixas S, Marques PI. Known Mutations at the Cause of Alpha-1 Antitrypsin Deficiency an Updated Overview of *SERPINA1* Variation Spectrum. Appl Clin Genet. 2021;14:173-194.

18. Patel D, Teckman JH. Alpha-1-Antitrypsin Deficiency Liver Disease. Clin Liver Dis 2018;22:643-655.

19. Ottaviani S, Barzon V, Buxens A, Gorrini M, Larruskain A, El Hamss R, Balderacchi AM, Corsico AG, Ferrarotti I. Molecular diagnosis of alpha1-antitrypsin deficiency: A new method based on Luminex technology. J Clin Lab Anal 2020;34:e23279.

20. Franciosi AN, Carroll TP, McElvaney NG. Pitfalls and caveats in α1-antitrypsin deficiency testing: a guide for clinicians. Lancet Respir Med. 2019 ;7:1059-1067

21. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR. A method and server for predicting damaging missense mutations. Nat Methods. 2010;7:248-9.

22. Ioannidis NM, Rothstein JH, Pejaver V, Middha S, McDonnell SK, Baheti S, Musolf A, Li Q, Holzinger E, Karyadi D, Cannon-Albright LA, Teerlink CC, Stanford JL, Isaacs WB, Xu J, Cooney KA, Lange EM, Schleutker J, Carpten JD, Powell IJ, Cussenot O, Cancel-Tassin G, Giles GG, MacInnis RJ, Maier C, Hsieh CL, Wiklund F, Catalona WJ, Foulkes WD, Mandal D, Eeles RA, Kote-Jarai Z, Bustamante CD, Schaid DJ, Hastie T, Ostrander EA, Bailey-Wilson JE, Radivojac P, Thibodeau SN,

Whittemore AS, Sieh W. REVEL: An Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants. Am J Hum Genet 2016;99:877-885.

23. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 1997;25:3389-402.

24. Sievers F, Higgins DG. Clustal omega. Curr Protoc Bioinformatics. 2014 ;48:3.13.1-16.

25. Crooks GE, Hon G, Chandonia JM, Brenner SE. WebLogo: a sequence logo generator. Genome Res. 2004;14:1188-90.

26. Medicina D, Montani N, Fra AM, Tiberio L, Corda L, Miranda E, Pezzini A, Bonetti F, Ingrassia R, Scabini R, Facchetti F, Schiaffonati L. Molecular characterization of the new defective P(brescia) alpha1-antitrypsin allele. Hum Mutat. 2009;30:E771-81.

27. Ronzoni R, Berardelli R, Medicina D, Sitia R, Gooptu B, Fra AM. Aberrant disulphide bonding contributes to the ER retention of alpha1-antitrypsin deficiency variants. Hum Mol Genet 2016;25:642-50.

28. Fra A, D'Acunto E, Laffranchi M, Miranda E. Cellular Models for the Serpinopathies. Methods Mol Biol 2018;1826:109-121.

29. Miranda E, Pérez J, Ekeowa UI, Hadzic N, Kalsheker N, Gooptu B, Portmann B, Belorgey D, Hill M, Chambers S, Teckman J, Alexander GJ, Marciniak SJ, Lomas DA. A novel monoclonal antibody to characterize pathogenic polymers in liver disease associated with alpha1-antitrypsin deficiency. Hepatology 2010;52:1078-88.

30. Dolmer K, Gettins PG. How the serpin α1-proteinase inhibitor folds. J Biol Chem. 2012 ;287:12425-32.

31. Laffranchi M, Berardelli R, Ronzoni R, Lomas DA, Fra A. Heteropolymerization of α-1-antitrypsin mutants in cell models mimicking heterozygosity. Hum Mol Genet. 2018;27:1785-1793..

32. Laffranchi M, Elliston EL, Miranda E, Perez J, Ronzoni R, Jagger AM, Heyer-Chauhan N, Brantly ML, Fra A, Lomas DA, Irving JA. Intrahepatic heteropolymerization of M and Z alpha-1-antitrypsin. JCI Insight. 2020;5:e135459.

33. Sifers RN, Finegold MJ, Woo SL. Alpha-1-antitrypsin deficiency: accumulation or degradation of mutant variants within the hepatic endoplasmic reticulum. Am J Respir Cell Mol Biol. 1989;1:341-5.

34. Huang X, Zheng Y, Zhang F, Wei Z, Wang Y, Carrell RW, Read RJ, Chen GQ, Zhou A. Molecular Mechanism of Z α1-Antitrypsin Deficiency. J Biol Chem. 2016 Jul 22;291:15674-86.

35. Jagger AM, Waudby CA, Irving JA, Christodoulou J, Lomas DA. High-resolution ex vivo NMR spectroscopy of human Z $\alpha_1$-antitrypsin. Nat Commun. 2020;11:6371.

36. Lomas DA, Irving JA, Arico-Muendel C, Belyanskaya S, Brewster A, Brown M, Chung CW, Dave H, Denis A, Dodic N, Dossang A, Eddershaw P, Klimaszewska D, Haq I, Holmes DS, Hutchinson JP, Jagger AM, Jakhria T, Jigorel E, Liddle J, Lind K, Marciniak SJ, Messer J, Neu M, Olszewski A, Ordonez A, Ronzoni R, Rowedder J, Rüdiger M, Skinner S, Smith KJ, Terry R, Trottet L, Uings I, Wilson S, Zhu Z, Pearce AC. Development of a small molecule that corrects misfolding and increases secretion of Z $\alpha_1$ -antitrypsin. EMBO Mol Med. 2021;13:e13167.

37. Ronzoni R, Heyer-Chauhan N, Fra A, Pearce AC, Rüdiger M, Miranda E, Irving JA, Lomas DA. The molecular species responsible for $\alpha_1$ -antitrypsin deficiency are suppressed by a small molecule chaperone. FEBS J. 2021;288:2222-2237.

38. Ronzoni R, Ferrarotti I, D'Acunto E, Balderacchi AM, Ottaviani S, Lomas DA, Irving JA, Miranda E, Fra A. The Importance of N186 in the Alpha-1-Antitrypsin Shutter Region Is Revealed by the Novel Bologna Deficiency Variant. Int J Mol Sci. 2021;22:5668.

39. Chambers JE, Zubkov N, Kubánková M, Nixon-Abell J, Mela I, Abreu S, Schwiening M, Lavarda G, López-Duarte I, Dickens JA, Torres T, Kaminski CF, Holt LJ, Avezov E, Huntington JA, George-Hyslop PS, Kuimova MK, Marciniak SJ. Z-$\alpha_1$-antitrypsin polymers impose molecular filtration in the endoplasmic reticulum after undergoing phase transition to a solid state. Sci Adv. 2022;8:eabm2094.

40. Barzon V, Ottaviani S, Balderacchi AM, Corino A, Piloni D, Accordino G, Coretti M, Mariani F, Corsico AG, Ferrarotti I. Alpha1-antitrypsin deficiency and M-like variants: how to improve the laboratory diagnosis of Mwurzburg and Mwhitstable alleles. Int J Med Sci 2022;23:9859.

41. Rizzo JM, Buck MJ. Key principles and clinical applications of "next-generation" DNA sequencing. Cancer Prev Res (Phila) 2012;5:887-900.

42. Balderacchi AM, Barzon V, Ottaviani S, Corino A, Zorzetto M, Wencker M, Corsico AG, Ferrarotti I. Comparison of different algorithms in laboratory diagnosis of alpha1-antitrypsin deficiency. Clin Chem Lab Med. 2021;59:1384-1391.

43. Fregonese L, Stolk J, Frants RR, Veldhuisen B. Alpha-1 antitrypsin Null mutations and severity of emphysema. Respir Med. 2008;102:876-84.

44. Mosella M, Accardo M, Molino A, Maniscalco M, Zamparelli AS. Description of a new rare alpha-1 antitrypsin mutation in Naples (Italy): PI*M S-Napoli. Ann Thorac Med. 2018 r;13):59-61.

45. Carpagnano GE, Santacroce R, Palmiotti GA, Leccese A, Giuffreda E, Margaglione M, Foschino Barbaro MP, Aliberti S, Lacedonia D. A New SERPINA-1 Missense Mutation Associated with Alpha-1 Antitrypsin Deficiency and Bronchiectasis. Lung. 2017;195:679-682

46. Matamala N, Lara B, Gomez-Mariano G, Martínez S, Retana D, Fernandez T, Silvestre RA, Belmonte I, Rodriguez-Frias F, Vilar M, Sáez R, Iturbe I, Castillo S, Molina-Molina M, Texido A, Tirado-Conde G, Lopez-Campos JL, Posada M, Blanco I, Janciauskiene S, Martinez-Delgado B. Characterization of Novel Missense Variants of SERPINA1 Gene Causing Alpha-1 Antitrypsin Deficiency. Am J Respir Cell Mol Biol. 2018;58:706-716.

47. Stein PE, Carrell RW. What do dysfunctional serpins tell us about molecular mobility and disease? Nat Struct Biol. 1995;2:96-113.
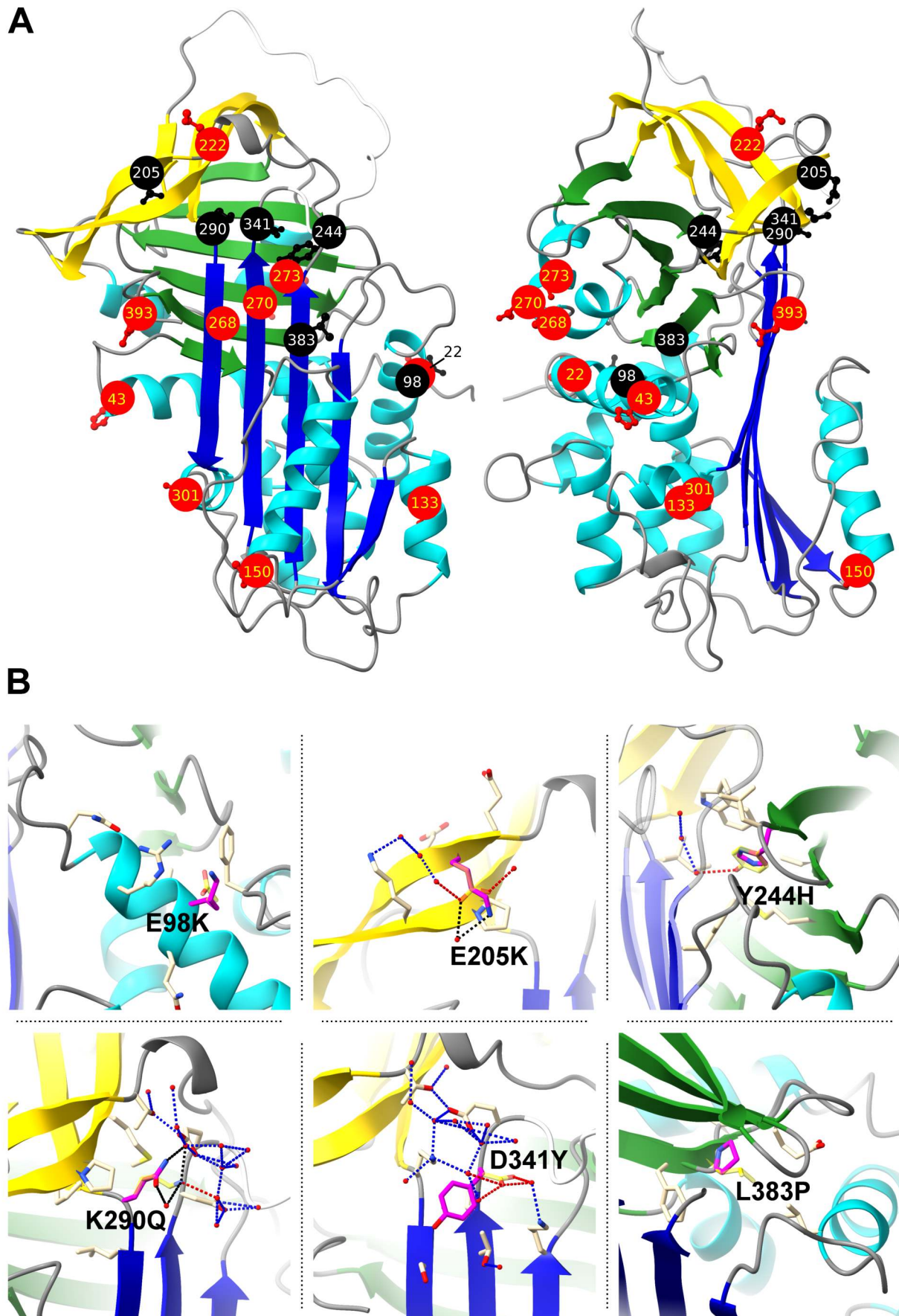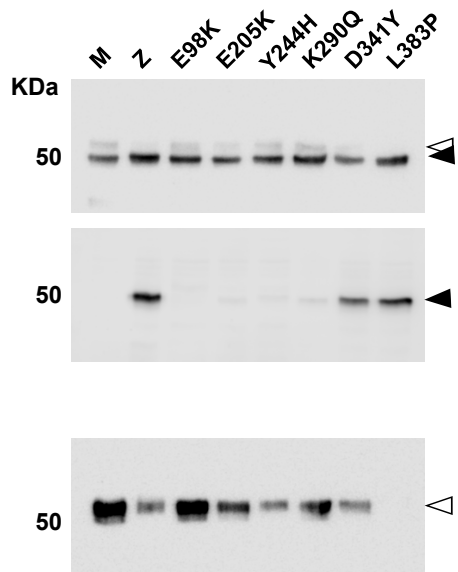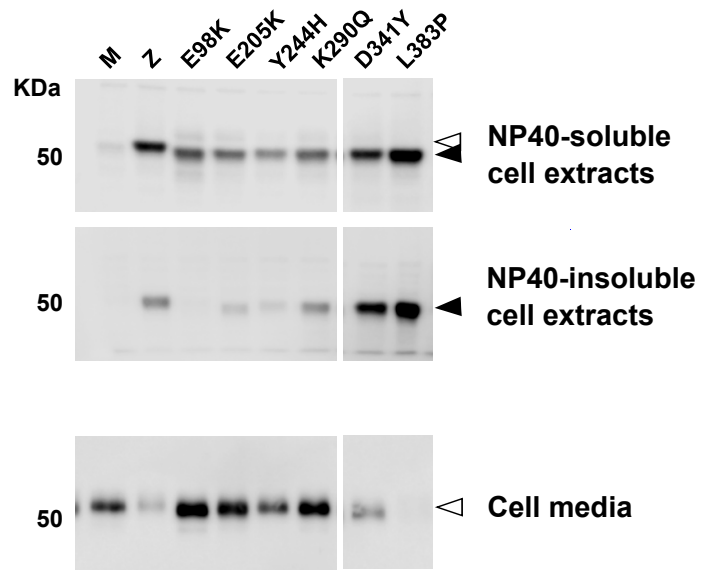
Figure 1

**A**

KDa

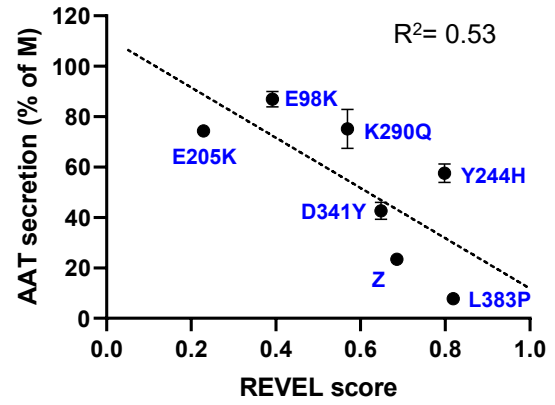M  Z  E98K  E205K  Y244H  K290Q  D341Y  L383P

50 ◁◀

50 ◀

50 ◁

**B**

KDa

M  Z  E98K  E205K  Y244H  K290Q  D341Y  L383P

50 ◁◀ **NP40-soluble cell extracts**

50 ◀ **NP40-insoluble cell extracts**

50 ◁ **Cell media**

Figure 2

Figure 3

**A**



**B**

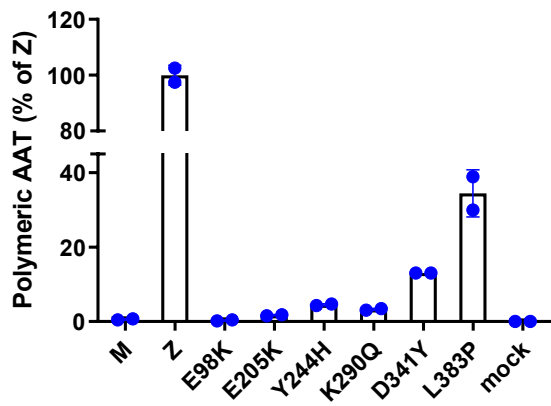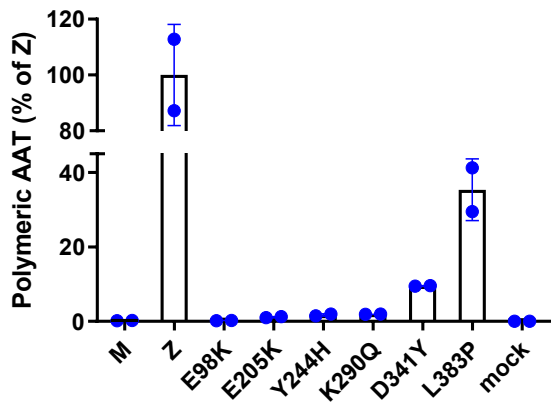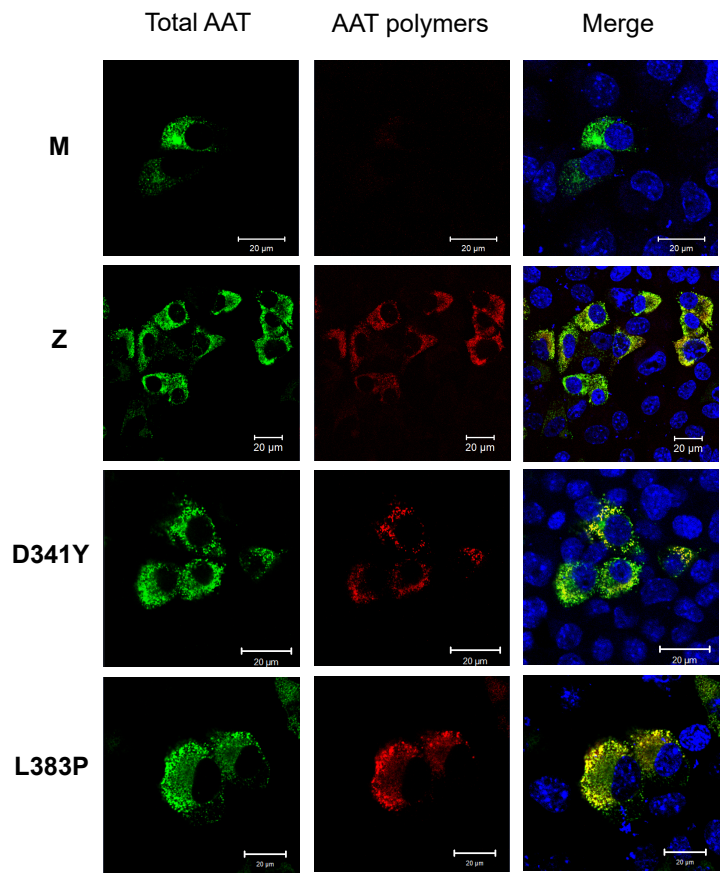|  | Total AAT | AAT polymers | Merge |
|---|---|---|---|
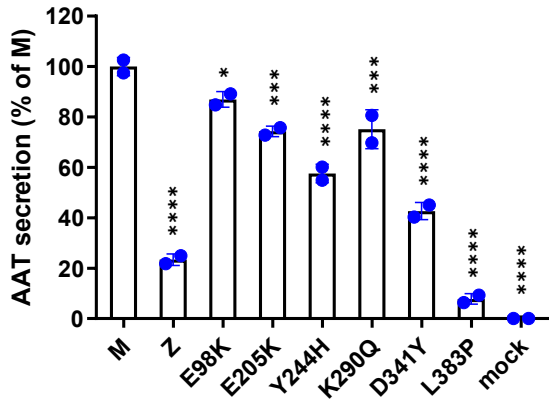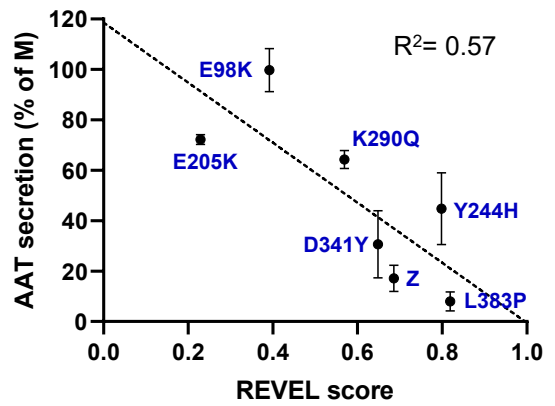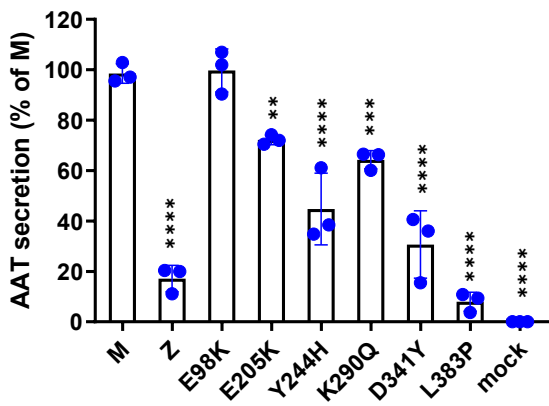


Figure 4

**Comprehensive clinical diagnostic pipelines reveal new variants in alpha-1-antitrypsin deficiency**

Stefania Ottaviani, Giulia Bartoli, Tomás P. Carroll, Fabrizio Gangemi, Alice M. Balderacchi, Valentina Barzon, Alessandra Corino, Davide Piloni, Noel G. McElvaney, Angelo G. Corsico, James A. Irving, Annamaria Fra, Ilaria Ferrarotti

**ONLINE DATA SUPPLEMENT**

## MATERIALS AND METHODS

### Biochemical and molecular diagnosis of AATD

Determination of AAT serum levels was performed by a rate immune nephelometric method (Immage 800, Beckman Coulter). The phenotype was determined by isoelectric focusing with immunofixation (Hydragel 18 A1AT IEF isofocusing kit, Sebia). DNA was isolated from whole peripheral blood, dried blood spot (DBS) samples or buccal swabs using a commercial extraction kit (QIAamp DNA Mini Kit or Investigator Kit (QIAGEN). Genotyping for the detection of S and Z variants was done by PCR with fluorescently labelled Taq-Man probes (VIC or FAM labels) on a LightCycler480 (Roche Diagnostics). We also evaluated the possible presence of 14 deficiency variants of the *SERPINA1* gene using the A1AT Genotyping Test (A1AT GT) (Progenika, Grifols) (1). The new mutations were identified by sequencing exonic (II-V) and part of the intronic (Ic and IV) portions, as well as flanking regions, of the *SERPINA1* gene by Sanger methods using the CEQ 8800 genetic analysis System (Beckman Coulter). When necessary, we performed Next Generation Sequencing (NGS) experiments to cover all the *SERPINA1* gene using the Illumina sequencing technology (iSeq100, Illumina).

### Cell transfection and lysis

HEK 293T/17 (ATCC CRL-11268) and Hepa 1-6 (ATCC CRL-1830) cell lines were cultured in DMEM/10%FBS and transiently transfected by polyethylenimine (PEI), as previously described (2,3). For each transfection, cells were grown to 95% confluence in 4 cm$^2$ wells. Plasmid-PEI complexes were prepared with 12 μg PEI and 2 μg plasmid, incubated for 20 min in 300 μL serum-free DMEM. The culture medium of the cells was replaced with 400 μl of complete medium, supplemented with the PEI-DNA complexes and the cells incubated for 5 h at 37°C. Transfected cells were then washed with Opti-MEM (Thermo Fisher Scientific) and further incubated for 20 h at 37°C with 700 μl of Opti-MEM. The cell culture media were then collected, centrifuged at 800 g for 5 min at 4°C to remove cell debris, while the cells were lysed in 250 μl of lysis buffer containing 1% NP40. NP40-soluble fractions were separated by centrifugation from the insoluble fractions, which were resuspended in equal volumes of lysis buffer and solubilized by sonication.

### Quantification of total AAT and polymeric AAT by ELISA

For quantification of total AAT levels in cell culture media, 96-well high-binding half-volume plates (Corning), coated overnight with 2 μg/ml rabbit polyclonal anti-AAT (Dako), were saturated with blocking buffer (PBS,0.25% BSA,0.05%Tween20) and then incubated for 1 h

at 37°C with serial dilutions of the standard plasma purified AAT (Athens Research & Technology) and of the samples in PBS/0.1% BSA. Bound AAT was detected with sheep anti-AAT-HRP (Abcam) and revealed with the TMB substrate (Merck). The reaction was terminated by 3M HCl and the absorbance at 450 nm measured by an ELISA plate reader. A similar ELISA was performed to quantify polymeric AAT, but coating of the plates was performed with 2 µg/ml of purified anti-AAT-polymer 2C1 mAb (4). The polymeric AAT standard was obtained by incubating purified plasma AAT for 40h at 56°C.

**Molecular dynamics**

Molecular dynamics (MD) simulations were performed by means of package GROMACS (5,6). The force field amber99-sb was used, with the PME method for the Coulomb interactions and a Lennard-Jones potential with a cut-off of 10 Å for the short-range interactions. The initial structure was completed by addition of hydrogens and solvated with TIP3P water in a simulation box with a minimum distance of 10 Å between solute and box boundaries. $Na^+$ and $Cl^-$ ions were added to reproduce a salt concentration of 150 mM and to neutralize the system. All simulations were conducted at 310 K. The system was first subjected to energy minimization, then equilibration at constant volume for 100 ps, and at constant pressure for 100 ps was performed before the production run at constant temperature and pressure. The temperature was kept constant by velocity rescaling with a characteristic time of 0.1 ps. The pressure was controlled using the Parrinello-Rahman method with a time constant of 1 ps and a compressibility of $4.5 \cdot 10^{-5}$ $bar^{-1}$. All simulations were conducted at 310 K. Structural analysis was performed by means of GROMACS utilities and in-house scripts. Visualisation with VMD 1.9.4 software (7).

**Table S1 –** List of reagents used in this study.

| Product | Manufacturer | Catalogue number |
|---|---|---|
| pcDNA3.1/Zeo (+) | Thermo Fisher | V86020 |
| QuikChangeII kit | Agilent | 200523 |
| GenEluteHP miniprep kit | Sigma | NA0150 |
| DMEM High glucose | Sigma | D5796 |
| FBS, Fetal bovine serum | Sigma | F7524 |
| NP40_NonidetP40 (IGEPAL) | Sigma | CA-630 |
| Protease inhibitors | Sigma | P8340 |

| | | |
|---|---|---|
| 4x SDS sample buffer | Biorad | 161-0747 |
| DTT, dithiotreitol | Sigma | D9779 |
| PVDF | Cytiva | 10600023 |
| Anti-human AAT polyclonal antibody | Dako, Agilent | A0012 |
| HRP-conjugated anti-rabbit IgG antibody | Cytiva | NA934V |
| ECL Plus | Euroclone | EMP011005 |
| 96-well high binding half-volume ELISA plates | Corning | 3690 |
| BSA, bovine serum albumin | Sigma | A7906 |
| Tween-20 | Sigma | P1379 |
| Sheep anti-AAT-HRP | Abcam | AMAB8768 |
| TMB | Sigma | T0440 |
| AlexaFluor 594-conjugated anti-mouse antibody | Thermo | A11032 |
| AlexaFluor 488- conjugated anti-rabbit antibody | Thermo | A11034 |
| DAPI_ NucBlue Fixed Cell Stain ReadyProbes reagent | Thermo | R37606 |
| ProLong antifade mounting medium | Thermo | P10144 |

**Table S2**- Oligonucleotide primers used for site-directed mutagenesis of expression vectors.

| AAT variant | Oligonucleotide sequences |
|---|---|
| **E98K (p.E122K)** | 5'-catgaaggcttccagaaactcctccgtaccc<br>5'-gggtacggaggagtttctggaagccttcatg |
| **E205K (p.E229K)** | 5'-gtcaaggacaccgagaaagaggacttccacg<br>5'-cgtggaagtcctctttctcggtgtccttgac |
| **Y244H (p.Y268H)** | 5'-agctgggtgctgctgatgaaacacctgggcaat<br>5'-attgcccaggtgtttcatcagcagcacccagct |
| **K290Q (p.K314Q)** | 5'-agcttacatttaccccaactgtccattactggaac<br>5'-gttccagtaatggacagttggggtaaatgtaagct |
| **D341Y (p.D341Y)** | 5'-ggctgtgctgaccatctatgagaaagggactgaag<br>5'-cttcagtccctttctcatagatggtcagcacagcc |
| **L383P (p.L407P)** | 5'-aaataccaagtctcccccccttcatgggaaaagtgg<br>5'-ccacttttcccatgaagggggggagacttggtattt |

**RESULTS**

**Novel *SERPINA1* allelic variants**

*Null alleles*

*c.-5+1800_1804delC (Q0$_{siracusa}$)*

The new mutation is located in the 5'UTR in exon 1C and consists of a C deletion at position c.-5+1800_1804. The deletion was first found in a 71-year-old male, never smoker, displaying intermediate AAT plasma levels (88 mg/dL) and suffering from dyspnoea. Subsequently we identified the mutation in a 57-year-old female smoker (20.5 pack/year), suffering from emphysema and COPD and displaying intermediate AAT levels (81 mg/dl) and in a 57-year-old male heavy smoker (45 pack/year) with emphysema and intermediate AAT plasma concentration (92 mg/dL). None of these individuals were immediate family members. Through the family screening of the third case, we found the mutation in a further 5 individuals: 3 brothers, a daughter and a son.

*P326Ter (Q0$_{sansiro}$)*

The variant was named Q0$_{sansiro}$ from the birthplace in Milan of the proband, since Q0$_{milano}$ already exists (8). It is caused by a C deletion at codon 326-327 that generates a stop codon (TGA). We found this mutation in *trans* with the rare deficient allele M$_{wurzburg}$ in a 33-year-old healthy male, referred because of low AAT plasma levels (39 mg/dL). The IEF displayed an M$_{wurzburg}$-like banding pattern (9). The proband, a never smoker, was apparently healthy but declined to undergo further medical investigation.

*K368Ter (Q0$_{firenze}$)*

The variant is due to an A>T transversion at codon 368 that produces a stop codon. This mutation, that we named Q0$_{firenze}$, was found in 2 related subjects in association with the normal M allele. The probands were a 49-year-old healthy male and his 23-year-old healthy son. They were referred because of intermediate AAT plasma levels (79 and 76 mg/dL, respectively).

*P369Pfs*5 (Q0$_{foggia}$)*

The mutation was detected in the proband, a 46 year old female never smoker suffering from chronic bronchitis and asthma, and in her 76 year old father. Both subjects were heterozygous for the S allele on genotyping but, interestingly, their phenotypes on IEF produced only the M banding pattern. The AAT concentration in plasma was below normal (81 and 69 mg/dL, respectively). We therefore performed sequence analysis that revealed, in both subjects, a C deletion at codon 369. This frameshift mutation created a stop TAA signal at codon 373. Family segregation and protein phenotypes suggested that the $Q0_{foggia}$ allele arose on an S allele background, a rare occurrence that makes the precise identification of each pathological mutation more challenging as it requires parental testing and *SERPINA1* sequencing to confirm.

*Missense alleles*

*T22A ($M_{asti}$)*

The proband was a 57-year-old male, never smoker, suffering from bronchitis and bronchiectasis. AAT plasma levels were at the lower end of the normal range (102 mg/dL). By sequencing we found a point mutation at codon 22 (A>G transition), changing an ACC (threonine) codon into a GCC (alanine) one (rs1243707738). The patient was also heterozygous for the rare $M2_{obernburg}$ (A284S) allele (10). Phenotyping on IEF did not reveal specific banding other than M, therefore we called the variant $M_{asti}$, from the birthplace of the proband.

*H43Q ($M_{napoli}$)*

This new mutation consists of a C>G transversion at codon 43 changing a CAC (histidine) codon into a CAG (glutamine) one (rs140744031). The probands were 2 unrelated patients: i) a 79 year old male, ex-smoker (pack/year not reported) with AAT serum concentration lower than normal (91 mg/dL), suffering from bronchitis, emphysema and dyspnea; ii) a 56 year old female, heavy smoker (40 pack/year), suffering from emphysema and displaying AAT serum concentration slightly lower than normal (114 mg/dL). Since phenotyping on IEF did not reveal specific banding other than M in both patients, we called the variant $M_{napoli,}$ from the birthplace of the first identified case.

*E98K ($P_{darfo}$)*

The proband was an Italian 83-year-old male suffering from vascular disease. Since the IEF pattern highlighted the presence of a band resembling a P variant, sequencing was performed, thus revealing a point mutation at codon 98 (G>A transition), changing a GAA (glutamic acid) codon into a AAA (lysine). The new mutation appeared in heterozygosity with the M1 allele. AAT plasma concentration was higher than normal (250 mg/dL), suggesting an acute phase response, which was confirmed by the elevated CRP level (2.57 mg/dL, normal range <0.8 mg/dL).

*D133N (P$_{savona}$)*

The variant consists of a G>A transition at codon 133 changing the GAT (aspartic acid) codon into AAT (asparagine). The proband was a 46-old-year male who suffers from haemophilia. He had high levels of AAT (172 mg/dl) but CRP levels were not available. The variant displayed a P-like IEF pattern.

*T150A (M$_{monza}$)*

The proband was a 45 year old male patient with AAT plasma levels slightly lower than normal (118 mg/dL). He was a heavy smoker (42 pack/year) suffering from emphysema and dyspnea. The new variant is due to an A>G transition (Thr/ACC to Ala/GCC) at codon 150 (rs1450215301). The IEF pattern was M-like, therefore we called the variant M$_{monza}$, from the birthplace of the patient.

*E205K (X$_{sarezzo}$)*

The variant is due to a G>A transition at codon 205 (Glu/GAA > Lys/AAA) and was found in 5 subjects living in the central-eastern Italian Alpine valleys, always in association with the normal M allele. Initially, it was detected in two unrelated subjects, a 42 year old male and a 64 year old female, during local blood donor screening (11). They both displayed AAT plasma levels at the lower limits of the norm (105 and 106 mg/dL); IEF data were not available for these samples. Then, the novel allele was detected in: i) a 14 year old male with normal AAT plasma concentration, and X-like IEF pattern; ii) an asthmatic 69 year old female and her 40 year old healthy daughter. Both were never smokers and displayed AAT plasma levels higher than normal with an X-like IEF pattern.

## K222T (F$_{milano}$)

The new mutation, an A>C transversion (Lys/AAG to Thr/ACG) at codon 222, was found in 3 unrelated patients: i) 48 year old female never smoker suffering from asthma and emphysema, displaying high level of AAT (194 mg/dL) with normal CRP concentration; ii) 53 year old male smoker with emphysema and normal AAT serum concentration (127 mg/dL); iii) a 50 year old female heavy smoker (40 pack/year) with lower AAT plasma levels (78 mg/dL). They all displayed an F-like IEF pattern, therefore we called the variant F$_{milano}$, from the birthplace of the first patient identified. An instance of the K222T (rs864622054) variant has previously been reported in ClinVar by a US-based commercial diagnostic laboratory without additional data.

## S236F (L$_{mestre}$)

The L$_{mestre}$ variant was found in a 52 year old female displaying L-like IEF pattern and high levels of AAT (224 mg/dL with CRP levels 1.22 mg/dL). The mutation consists of a C>T transition at codon 236 changing a TCC (serine) into a TTC (phenylalanine).

## Y244H (P$_{dublin}$)

The proband was a 60 year old Irish male displaying a high AAT plasma concentration (245 mg/dL, subsequently 181 mg/dL but CRP measurements were not available). Liver enzymes were elevated but ultrasound was normal, with cough and wheeze reported as respiratory symptoms. IEF revealed a P-like banding pattern in combination with M. The mutation was found in exon 3 at codon 244 and consists of a T>C transition changing TAC (tyrosine) into CAC (histidine).

## T268N (M$_{andria}$)

The M$_{andria}$ variant was found in a 50 year old female, occasional smoker (6.7 pack/year), reporting asthma, bronchitis and emphysema. The mutation consists of a C>A transversion at codon 268 changing ACC (threonine) to AAC (asparagine). The patient was also heterozygous, likely in *trans*, for the non-pathological X$_{christchurch}$ variant. The AAT plasma level was lower than normal (88 mg/dL) and IEF showed a M1/X pattern.

## D270N (L$_{bressanone}$)

The proband was an 18 year old female referred because her phenotype was reported as ZZ during neonatal screening based on IEF (12). The patient was subsequently found to be heterozygous for the Z allele and displayed AAT plasma levels lower than normal (97 mg/dL with a CRP concentration of 2.71 mg/dL). The mutation is reported in ClinVar (rs772436715) by a US-based company with no other information provided. Through family screening we identified the $L_{bressanone}$ allele in the mother of the proband, thus confirming that $L_{bressanone}$ and Z alleles occur in *trans*.

### T273N ($M_{ancona}$)

The mutation was identified in a Z heterozygous 63 year old female with intermediate AAT plasma concentration (66 md/dL) and MZ IEF pattern. The $M_{ancona}$ variant consists of a C>A transversion, changing a threonine into an asparagine at codon 273. It has been reported in ClinVar (rs752776707) by a US-based commercial laboratory without further data.

### K290Q ($C_{tiberias}$)

DNA sequencing revealed a heterozygous mutation at codon 290 in a 59 year old Israeli male ex-smoker (31 pack/year), suffering from dyspnea and displaying abnormal liver enzymes. An A>C transversion in the exon 4 causes a change from lysine to glutamine. AAT plasma levels were normal (137 mg/dL). The same mutation was then found in an 80 year old Italian woman suffering from bronchiectasis, displaying AAT plasma levels at the lower limits of normal (105 mg/dL). IEF pattern resembled a C variant.

### S301R ($V_{verceia}$)

The $V_{verceia}$ mutation was identified by sequencing the DNA sample from a 50 year old healthy female non-smoker, that presented with normal AAT plasma levels (132.6 mg/dL) and an IEF pattern with bands resembling the V variant. The $V_{verceia}$ allele is caused by a C>G transversion at codon 301 changing a serine (AGC) into an arginine (AGG).

### D341Y ($S_{milano}$)

The proband was a 44 year old male never smoker suffering from bronchiectasis. AAT plasma concentration was at the lower limit of normal (119 mg/dL) and IEF pattern was S-like. DNA sequencing revealed heterozygosity for a G>T transversion at codon 341 (GAC>TAC) in exon

5. An equivalent mutation is reported in ClinVar (rs143370956) by a US-based commercial laboratory without further details. The more conservative P$_{saint\ albans}$ (D341N) mutation at this site has previously been observed to produce levels comparable to M in a cell expression model (13).

### L383P (M$_{campolongo}$)

The mutation, a T>C transition in exon 5, was found in a 58 year old male with AAT plasma concentration lower than normal (72 mg/dL with CRP levels 0.99 mg/dL) and M phenotype at IEF. The AAT novel variant is located at codon 383 and produces a leucine to proline change.

### Q393K (X$_{magenta}$)

The proband was a 56 year old male with normal AAT plasma level (121 mg/dL) and X-like IEF pattern. Family screening contributed to find the X$_{magenta}$ allele in the 2 sons of the proband; they both displayed normal AAT plasma concentration (123 and 157 mg/dL). The mutation consists in a glutamine (CAA) to lysine (AAA) change at codon 393.

**SUPPLEMENTAL FIGURES**

Figure S1. Detailed views of residues predicted to be minimally-perturbing are shown, in which the wild-type residue is shown in dark yellow stick, the mutation in magenta stick, nearby residues are pale yellow, and water molecules red spheres. Polar interactions present in both the wild-type and mutant are shown as dashed blue lines, those present only in the wild-type are red, and those only in the mutant are yellow; black dashes indicate an altered but preserved bond. Secondary structure elements are β-sheets A (blue), B (green) and C (yellow), with helices in cyan and loop regions grey.

Figure S2. Left: molecular dynamics analysis of Ramachandran angles of residues 181-192 for Y244H (red) and M (blue). Right: selected residues are highlighted in red on the structure, with residue H244 explicitly shown. Significant differences are found for Y244H with respect to M, indicating a structural change.

Figure S3. Average number of hydrogen bonds of residue 244 with other residues in the case of M and Y244H variant (see structure aside). The interactions of residue 244 are increased upon mutation, due to new bonds with Lys 193, Lys 243 and Gly 192.

Figure S4. Average number of hydrogen bonds between residue 341 and other residues in the case of wild type (M) and variant D341Y, as obtained from MD simulations. In this case there is a significant decrease in the total number of hydrogen bonds of D341Y with respect to M. In the picture on the right hand side of the table some of the possible hydrogen bonds are shown with blue dashed lines. The interaction with strand s6A in the case of M is based on bonds with Lys 191, Gly 192 (backbone), Lys 193, and Trp 194, collectively amounting to an average of 1.18, while in the D341Y variant its average is 0.74. The interactions in the hinge region are shown by the two bonds with Glu 342 (backbone) and Lys 343, which in M collectively amount to an average of 1.09, and are essentially absent in the case of the D341Y variant.

**REFERENCES**

1.  Ottaviani S, Barzon V, Buxens A, Gorrini M, Larruskain A, El Hamss R, Balderacchi AM, Corsico AG, Ferrarotti I. Molecular diagnosis of alpha1-antitrypsin deficiency: A new method based on Luminex technology. J Clin Lab Anal 2020;34:e23279.

2.  Ronzoni R, Berardelli R, Medicina D, Sitia R, Gooptu B, Fra AM. Aberrant disulphide bonding contributes to the ER retention of alpha1-antitrypsin deficiency variants. Hum Mol Genet 2016;25:642-50.

3.  Fra A, D'Acunto E, Laffranchi M, Miranda E. Cellular Models for the Serpinopathies. Methods Mol Biol 2018;1826:109-121.

4.  Miranda E, Pérez J, Ekeowa UI, Hadzic N, Kalsheker N, Gooptu B, Portmann B, Belorgey D, Hill M, Chambers S, Teckman J, Alexander GJ, Marciniak SJ, Lomas DA. A novel monoclonal antibody to characterize pathogenic polymers in liver disease associated with alpha1-antitrypsin deficiency. Hepatology 2010;52:1078-88.

5.  Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B, Lindahl E, GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers, SoftwareX 2015; 1–2:19-25,

6.  Hess B, Kutzner C, van der Spoel D, Lindahl E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. J Chem Theory Comput. 2008;4:435-447.

7.  Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. J Mol Graph. 1996;14(1):33-28.

8.  Rametta R, Nebbia G, Dongiovanni P, Farallo M, Fargion S, Valenti L. A novel alpha1-antitrypsin null variant (PiQ0Milano ). World J Hepatol 2013;5:458-61.

9.  Barzon V, Ottaviani S, Balderacchi AM, Corino A, Piloni D, Accordino G, Coretti M, Mariani F, Corsico AG, Ferrarotti I. Alpha1-antitrypsin deficiency and M-like variants: how to improve the laboratory diagnosis of Mwurzburg and Mwhitstable alleles. Int J Med Sci 2022;23:9859

10. Faber JP, Poller W, Weidinger S, Kirchgesser M, Schwaab R, Bidlingmaier F, Olek K. Identification and DNA sequence analysis of 15 new alpha 1-antitrypsin variants, including two PI*Q0 alleles and one deficient PI*M allele. Am J Hum Genet 1994;55:1113-21.

11. Ferrarotti I, Corda L, Gatta N, Ottaviani S, Braghini A, Arici M, Lanzani G, Benini F, Corsico A, Balbi B. Screening of alpha-1 antitrypsin deficiency in a blood donors cohort of the North-Italian area [abstract]. Eur Respir J 2015 46: PA1135.

12. Pittschieler K, Massi G. Alpha 1 antitrypsin deficiency in two population groups in north Italy. Padiatr Padol 1988;23:307-311.

13. Holmes MD, Brantly ML, Crystal RG. Molecular analysis of the heterogeneity among the P-family of alpha-1-antitrypsin alleles. Am Rev Respir Dis 1990;142:1185-92.
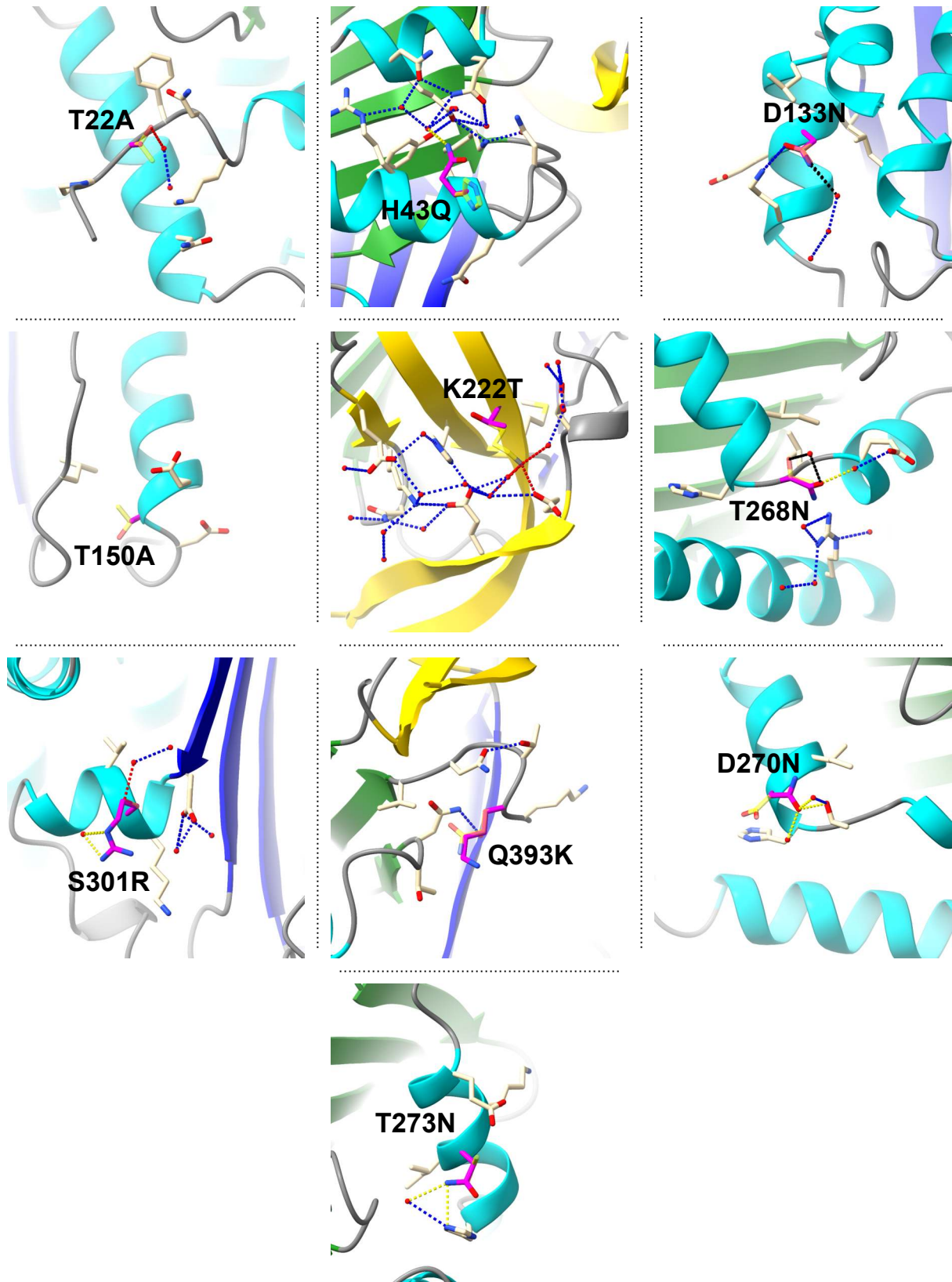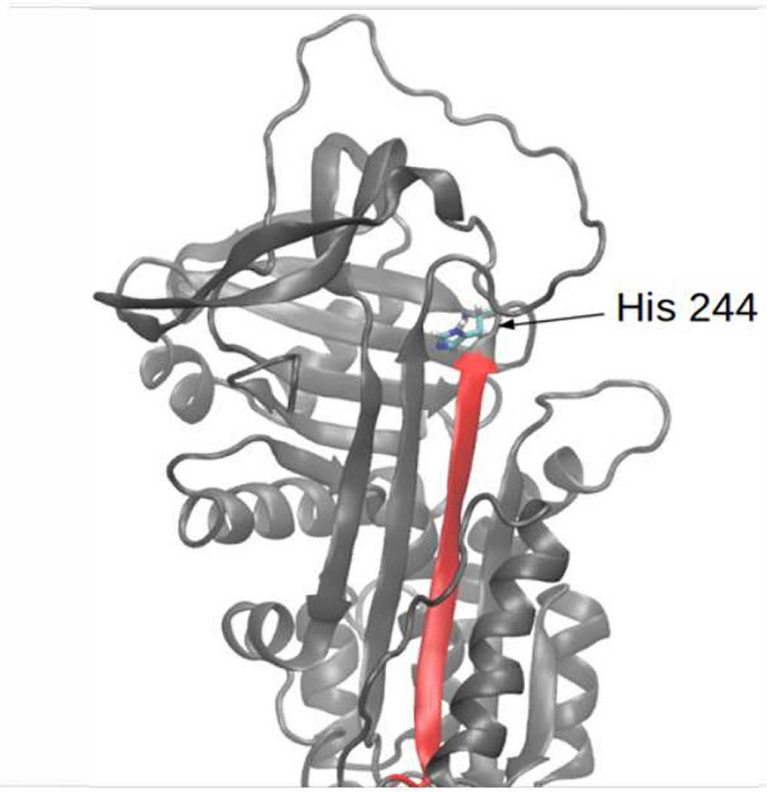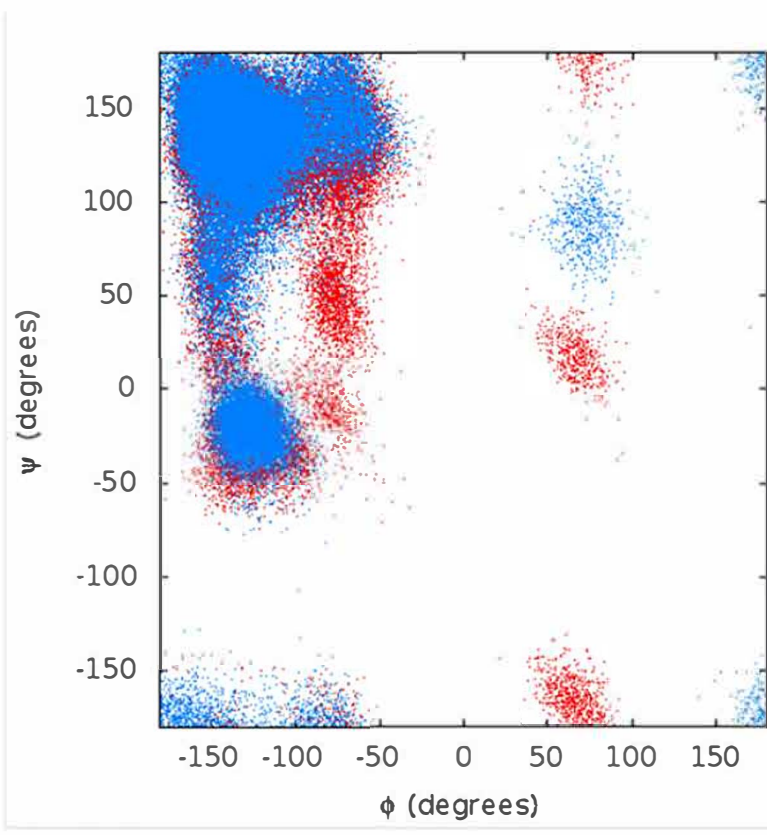
Figure S1

His 244

Figure S2

| Residue | M | Y244H |
|---------|------|-------|
| Gly 192 | 0.00 | 0.08 |
| Lys 193 | 0.00 | 0.38 |
| Lys 243 | 0.00 | 0.08 |
| Gly 246 | 0.08 | 0.05 |
| Ala 248 | 0.94 | 0.96 |
| Total | 1.03 | 1.55 |
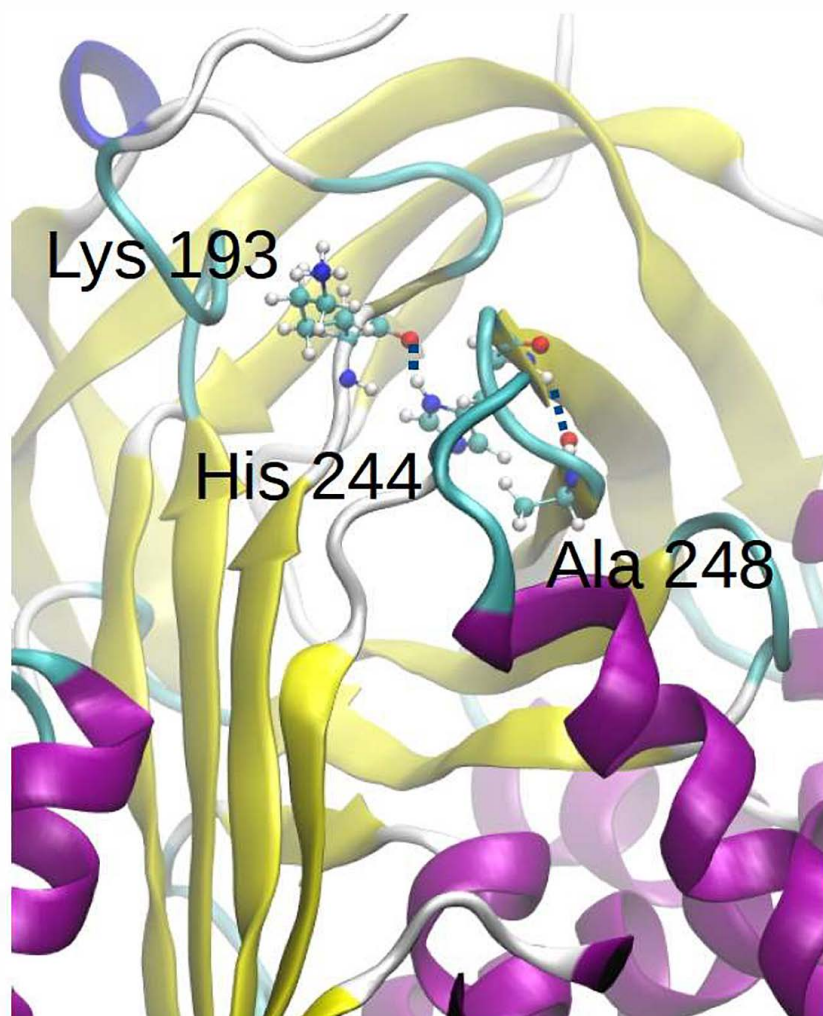


Figure S3

| Residue | M | D341Y |
|---------|------|-------|
| Asn 111 | 0.03 | 0.00 |
| Lys 191 | 0.19 | 0.02 |
| Gly 192 | 0.36 | 0.05 |
| Lys 193 | 0.23 | 0.00 |
| Trp 194 | 0.39 | 0.68 |
| Lys 290 | 0.07 | 0.00 |
| Ser 292 | 0.00 | 0.04 |
| Glu 342 | 0.21 | 0.00 |
| Lys 343 | 0.88 | 0.00 |
| Gly 344 | 0.06 | 0.04 |
| Total | 2.44 | 0.83 |



Figure S4