On the diversity of antithrombin proteoforms: the role of a diagnostic mass spectrometry-based test for antithrombin deficiency

Kruijt, M.

5

Novel insights into Antithrombin Deficiency enabled by Mass Spectrometry-based Precision Diagnostics

M. Kruijt[1], M.E. de la Morena-Barrio[2], J. Corral[2], C.M. Cobbaert[1], L.R. Ruhaak[1]

1) Department of Clinical Chemistry and Laboratory Medicine, Leiden University Medical Center, Leiden, The Netherlands

2)Servicio de Hematología, Hospital Universitario Morales Meseguer, Centro Regional de Hemodonación, Universidad de Murcia, IMIB-Pascual Parrilla, CIBERER-ISCIII, Murcia, Spain

Published in Journal of Thrombosis and Haemostasis (2024)

Abstract

Background

Although P5 (preventive, personalized, predictive, participatory, psychocognitive) medicine and patient focused healthcare are gaining ground in various healthcare areas, the diagnosis of antithrombin deficiency (ATD) is still based on crude diagnostic tests clustering patients into clinically heterogeneous subgroups whereby relevant thrombophilia phenotypes may go unnoticed. Due to limited clinical studies on harmful effects of high-risk ATD subtypes, generic treatments are still the norm.

Objectives

To unravel the heterogeneity of ATD, a mass spectrometry (LC-MRM-MS)-based test for antithrombin was developed allowing molecular characterization of the antithrombin proteoforms in patient plasma. This study provides the first insight into the tests' clinical performance.

Methods

Plasma from 91 unrelated ATD patients and 41 patients with a congenital disorder of glycosylation affecting antithrombin glycosylation were characterized functionally, genetically, and analyzed by LC-MRM-MS. An established data analysis strategy was applied for quantitation and molecular characterization of antithrombin proteoforms.

Results

The test recognized patients with a quantitative defect, discriminated between type I and type II ATD, and identified variant proteoforms. Overall, the diagnostic sensitivity for ATD was 100 % for LC-MRM-MS compared to 81.1 % by the functional test. Type II ATD, a subtype prone to misdiagnosis, revealed an even larger difference of 100 % identification by LC-MRM-MS versus 56.8 % by functional test.

Conclusions

The qualitative and quantitative MS-based AT-test can serve as a platform for investigating the molecular basis of the clinical heterogeneity of ATD. This Precision Diagnostics approach for ATD can lower diagnostic uncertainty and modernize the ATD diagnostic and clinical pathways.

Introduction

Antithrombin deficiency (ATD) is a clinically heterogenous disorder due to the large number of possible genetic and post-translational modifications influencing the expression, secretion, functionality, and stability of antithrombin (AT), a key coagulant serpin [49, 252-254]. Although AT was first mentioned in 1939, and the first notion of ATD was in 1965, current diagnostic tests still lack sensitivity and specificity leading to missed diagnoses and cannot distinguish patients at high and low risk for venous thromboembolism (VTE) [20, 68, 255]. Without anticoagulant treatment ATD leads to a high annual risk of recurrent venous thromboembolism (VTE) of 8.8 % [34]. Consequently, it is suggested to treat ATD patients indefinitely with an anticoagulant, such as vitamin K antagonist drugs or DOACs [149, 154, 256]. Such treatment brings along bleeding risk which, although seemingly low at an annual risk of 0.5 - 0.8 %, contributes to morbidity while the annual VTE recurrence risk is still 2.7 %. The current clinical care pathway for ATD, suffering from the unresolved clinical heterogeneity of ATD and a one-size-fits-all approach, leads to over- and undertreatment of individual ATD patients, and deserves modernization.

To reduce diagnostic uncertainty in the ATD clinical pathway and facilitate P5 medicine (preventive, personalized, predictive, participatory, psychocognitive), a better understanding of the spectrum of AT-proteoforms and their pathological diversity is key [257]. The current diagnostic tests for ATD do not meet the required clinical performance to distinguish between low and high risk ATD impeding personalized treatment. Commercially available AT activity and antigen tests only allow classification of patients into two subtypes, type I (quantitative) and type II (qualitative) ATD with varying clinical severity [133], and lack sensitivity for specific types of ATD, resulting in missed diagnoses [14, 62, 166]. Genetic testing offers specific insight into the exact molecular defect. However, it does not identify defects caused by post-translational modifications (PTMs), such as N-glycosylation, and only provides a blueprint of how the protein may be expressed. Importantly, N-glycosylation may have key functional and clinical relevance [62]. This highlights that clinical phenotypes are only indirectly caused by genes, and instead proteins, playing a direct role in diseases, are likely to provide additional information [242, 243]. Therefore, molecular information on the proteoforms present in ATD patients may hold the key to better identify and characterize ATD patients.

Molecularly, ATD is a complex disorder. Type I ATD is caused by genetic variants impeding the translation and/or secretion of the mutant proteoform resulting in a decrease in AT concentration of around 50 % (with the remaining AT originating from the wildtype allele) [258]. In contrast, type II ATD or ATD caused by aberrant glycosylation leads to the presence of variant AT proteoforms in the circulation with varying clinical severity and challenging diagnoses due to the unpredictable influence of the variant proteoforms on the detection by AT activity tests [14, 133, 142, 166]. Current diagnostic tests focus on the activity or concentration of the total AT pool (**Figure 1**) obtained through indirect approaches, e.g. based on chromogenic or latex agglutination tests. It is therefore not surprising that pleas have been made to introduce refined molecular tests for AT, aiming to fully characterize, quantitatively and qualitatively, AT proteoforms in plasma [142, 235, 259]. Ideally, the envisioned molecular test(s) should incorporate all molecular features of AT including glycosylation.

Recently, a test based on liquid-chromatography coupled to multiple-reaction-monitoring mass spectrometry (LC-MRM-MS) was developed [218] for the molecular characterization of plasma AT proteoforms. In contrast to current tests used in the diagnosis of ATD, MS enables direct monitoring of the AT proteoforms present in a patient sample allowing both quantitation (in µmol/L) and (in theory) molecular characterization (**Figure 1**) [205]. The test was analytically validated, which verified the tests' ability to quantify the AT concentration of clinical samples in clinically relevant ranges, and a case report indicated that the theorized molecular characterization is feasible [140]. However, to

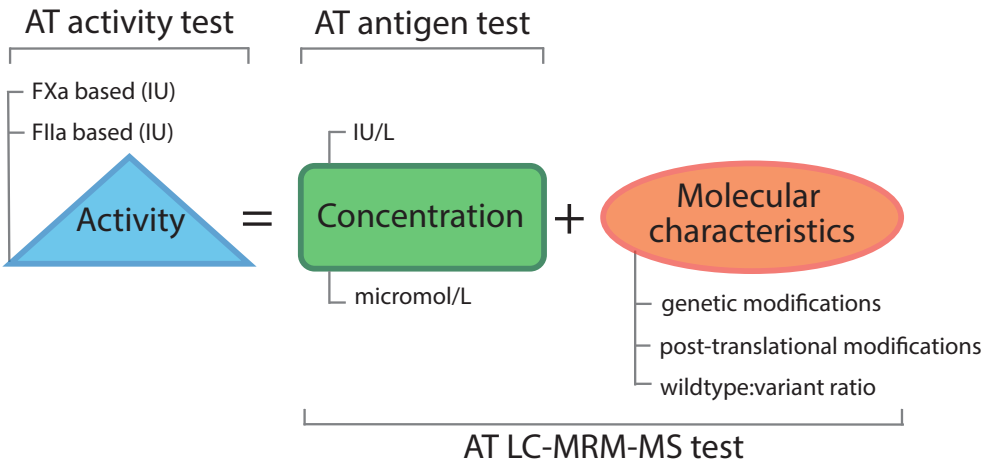


Figure 1. Focus of contemporary (activity and antigen test) and new tests (LC-MRM-MS) for ATD.

assess the potential of the LC-MRM-MS based AT-test for improving patient management, its clinical performance must be evaluated.

Clinical performance studies ideally require a prospective setup in which patients with suspected clinical symptoms for (in this case) ATD are included and tested with both the current and the comparator test [18]. Due to the low frequency of ATD in the VTE population in combination with the lack of experience on the ability of the test to identify mutations, it was decided to first assess the tests' scientific validity retrospectively in a hereditary ATD patient cohort. Moreover, ATD has also been reported in patients with type I congenital disorders of glycosylation (CDG), who suffer from a systemic N-glycosylation defect leading to a varying degree of AT hypoglycosylation. To assess the tests' ability to detect alterations in AT glycosylation a cohort of patients with CDG was also evaluated. Notably, the clinical care pathway for CDG, and therefore the clinical need is different in CDG patients compared to ATD patients. The aim of this study was to assess the ability of the test to identify ATD, stratify between type I and II ATD, and investigate to what extent the test enables full molecular characterization of the AT proteoforms present in clinical samples.

Methods

Patient samples

Samples from 132 unrelated patients with genetically confirmed ATD or CDG, of which the majority had a history of thrombosis, were selected for LC-MRM-MS analysis [166, 260]. Samples were processed within 24 hours of blood draw by centrifuging citrate tubes (2200 g, 20 min), collecting the platelet poor plasma and storing them at -80 °C. All patients gave informed consent following ethical guidelines, as approved by the institutional review board of Hospital Universitario Morales Meseguer and in accordance with the Declaration of Helsinki of 1964 and its subsequent amendments.

Samples from hereditary ATD patients (N = 91) were selected from a biobank of the Hospital Universitario Morales Meseguer, which contains over 350 unrelated cases recruited between 1998-2023. Patients underwent thrombophilia screening and were recruited via three main routes: 1) due to a clinical event (e.g. VTE, N = 77); 2) screening indicated by family studies (n = 9, note that only unrelated samples were analyzed for the study described here); 3) screening in a gynecological setting or when prescribed contraceptives (n = 5). Selection aimed to include a varying repertoire of mutations which were genetically confirmed to carry a pathogenic genetic variant in the AT gene *SERPINC1*.

Most patients enrolled in this study have maintained antithrombin deficiency and thrombosis, which are explained by *SERPINC1* variants. These cases are easily diagnosed by functional and genetic methods. However, we also identified cases with much more complex diagnosis, as they have transient antithrombin deficiency explained by different mechanisms, from structural constrains resulting in pathogenicity only under stress conditions, to the use of different functional methods that may not detect the functional consequences of the *SERPINC1* variant. The characteristics of these cases have been published [16]. Thus, we systematically sequenced *SERPINC1* gene in all 350 cases with suspicion of antithrombin deficiency recruited during more than 25 years, and selected samples of cases carrying different *SERPINC1* variants independently of having normal antithrombin levels. Twelve *SERPINC1* variants selected for this study caused transient antithrombin deficiency, and the samples evaluated in this study had no antithrombin deficiency caused by the variant (p.Val30Glu; p.Arg45Trp; p.Pro73Leu; p.Arg79His; p.Val137Ala; p.Arg177Cys; p.Gly199Arg; p.Asn224His; p.Glu227Lys; p.Ala416Ser, p.Ser426Leu and p.Pro439Thr).

To assess the potential of the LC-MRM-MS test to identify PTMs, patients with CDG (N = 41) were also included. CDG leads to systemic changes in the glycosylation, affecting AT as well as other proteins, and transferrin glycoforms were analyzed by HPLC to ascertain the diagnosis of CDG (which is the gold-standard for CDG diagnostics, although diagnosis is still challenging [261])).

For comparative purposes, samples from 37 healthy controls were also analyzed. These controls were recruited via the Leiden University Medical Center Voluntary Donor Service (as approved by the Leiden University Medical Center ethics board) which recruits (apparently) healthy hospital employees to donate blood for studies.

Genetic analysis

Genetic defects were identified by next-generation sequencing of the *SERPINC1* gene, or 72 genes potentially involved in CDG [60, 249]. For *SERPINC1* the exons, flanking regions and 1500 bp of the promotor region were amplified by polymerase chain reactions (PCR, Expand Long Template Polymerase, Roche, Spain) and sequenced with ABI Prism Big Dye Terminator v3.1 Cycle sequencing kit on a 3130xl Genetic Analyzer (Applied Biosystems, Spain). For detailed information, see de-la Morena-Barrio et al (2012) [262]. Obtained sequences were compared with a reference sequence (GenBank NG_012462.1) using SeqScape v2.5 Software (Applied Biosystems). Of note, gross deletions were analyzed by multiplex ligation-dependent probe amplification (MLPA) using the SALSA®MLPA® Kit P227 SERPINC1 (MRC-Holland, Amsterdam, the Netherlands) [250]. For CDG-patients, exons and flanking regions of the PMM2 gene were PCR amplified and sequenced. If no mutation was found, whole exome sequencing was performed on an Ion Proton™ platform (Life Technologies, Madrid, Spain) using the AmpliSeq™ kit and Ion Reporter™ Software. Results were validated using corresponding primers [60]. Most CDG patients (N = 37) were subtyped as PMM2-CDGs, with the remaining four patients carrying other subtypes (DPAGT1-CDG, MPI-CDG, ALG12-CDG and SSR4-CDG).

Antithrombin Activity

AT activity was measured in all samples except one (N = 131) using an in-house developed method based on a chromogenic anti-FXa method (HemosIL Liquid Antithrombin, #0020008900, Werfen) using S-2765 substrate and bovine Factor Xa read-out on a Synergy HT plate reader. The method was verified towards an automated coagulometer (ACL-TOP), using chromogenic substrate (#0020008910, Werfen) and bovine Factor Xa (#0020008920, Werfen). Calibration was performed by serial dilution of a reference plasma generated by pooling 100 healthy blood donors. Normal ranges, as tested in 250 healthy blood donors, are 80 - 120 %.

LC-MRM-MS test

The LC-MRM-MS test was developed and analytically validated according to established laboratory

guidelines, as described elsewhere [218]. Patient samples were analyzed in duplicate whereas controls were analyzed in duplicate (N = 11) or singlicate (N = 26). Upon suspicion of a variant proteoform in a sample the presence of the proteoform was confirmed as described elsewhere (Kruijt et al., manuscript submitted). A system suitability test sample was measured in fivefold before and after each experimental batch to monitor system performance. Furthermore, two quality control (QC) samples were included in threefold in each experimental batch to monitor the total test performance. The two QCs showed means ± CV of 1.37 µmol/L ± 6.0 % and 1.18 µmol/L ± 5.9 % (based on quantitative peptide LVSAN). Additionally, ion ratios were evaluated for the total data set to ensure data quality and exclude interferences.

Data analysis

Raw MS files were automatically interpreted using Agilent MassHunter Workstation Quantitative Analysis (v10.0), followed by manual inspection and adjustment if necessary. Results were imported as excel files into R studio (v1.4.1717). All data analysis steps (quality control, data verification and exploration and reporting) were executed by in-house built scripts apart from proteoform verification, which was examined manually. In depth information and examples of the data analysis process are found elsewhere (**Supplementary information** and Kruijt et al., manuscript submitted).

Results

The LC-MRM-MS test was applied in a cohort of genetically confirmed ATD patients (N = 91, see **Table 1** for patient details). The majority of ATD patients had clinical symptoms, although asymptomatic ATD patients were also included, originating from family studies or gynecological thrombophilia screenings. As PTMs are also known to affect AT function and are included in the LC-MRM-MS test, a cohort of CDG patients (N = 41) was also analyzed [62]. These patients suffer from a systemic defect in their N-glycosylation which, among other proteins, affects the glycosylation of AT. CDGs are diagnosed via a different diagnostic pathway than hereditary ATD, in general by a pediatrician, with multiple clinical effects in these patients. Thrombophilia is not a hallmark feature of CDG patients, although thrombosis occurs more often in CDG patients than in the general population and therefore frequent thrombophilia screening is recommended (e.g. yearly or before surgery/during pregnancy) [263, 264].

The LC-MRM-MS test quantifies the concentration of 23 AT peptides, from which the overall AT concentration (based on a quantifier peptide) as well as in-depth information on the status of each protein stretch (based on qualifier peptides) can be derived (Kruijt et al., manuscript submitted). The anticipated output of the LC-MRM-MS test per patient was: 1) the AT concentration in µmol/L and 2) the presence of molecular defects either caused by mutations or variation in the glycosylation. Together this would allow the identification of ATD and subtype classification.

Table 1. Summary characteristics of ATD cohorts studied by LC-MRM-MS.

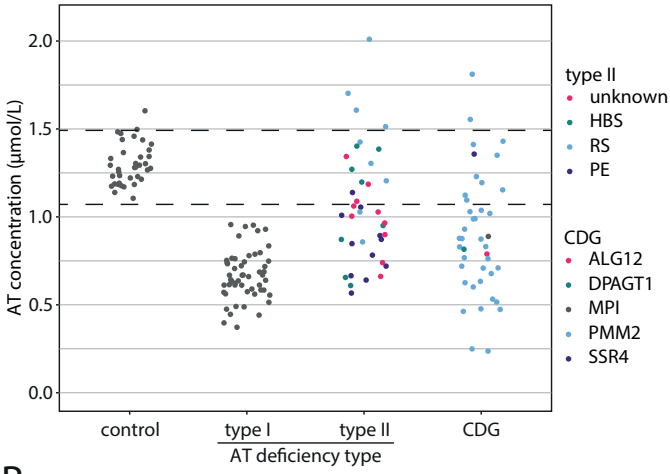
	Patients (N, %)	Activity (%)	Thrombosis reported (%)†
Hereditary ATD	91	33-102	
- Type I	53 (58.2)	33 – 88	90.6
- Type II	38 (41.8)	40 – 102	58.8*
- Subtype unknown	10 (11.0)	40 – 97	100*
- HBS	8 (8.8)	44 – 90	50.0
- RS	9 (9.9)	50 - 96	55.5
- PE	11 (12.1)	46 – 102	45.5
CDG	41	9 - 116	NA
Controls	37	ND	NA

ND not determined, NA not available. * 4 samples did not have a complete clinical background. † Thrombosis reported entails clinically confirmed DVT, PE, cerebral venous thrombosis or mesenteric vein thrombosis.

Antithrombin concentration

A quantitative deficiency resulting in low AT concentrations is a hallmark of ATD, specifically for type I ATD. To this end, the concentration of the total AT proteoform pool was analyzed by LC-MRM-MS. All control samples showed concentrations within the reference intervals (1.07 – 1.49 µmol/L), apart from two samples that were slightly higher than the reference interval (RI) (**Figure 2**). As expected, all patients carrying a type I ATD had a low AT concentration ranging between 0.37 – 0.96 µmol/L. Type II ATD, usually defined by low AT activity levels but normal antigen levels, showed more heterogeneous concentrations ranging between 0.57 - 2.01 µmol/L. Of the 38 type II ATD patients, 23 patients had an AT concentration below the RI and eleven patients had a concentration within the RI. Interestingly, four type II ATD patients presented with concentrations above the RI, all harboring a mutation in either Arginine-425 or Serine-426, the reactive site of AT. Samples from CDG patients showed a wide concentration range, between 0.24 to 1.81 µmol/L, with 30 patients presenting with low concentrations, nine with normal concentrations and two patients with elevated concentrations. These results highlight the variability in AT concentrations found in type II ATD and CDG patients. Taken together, 76/91 (83.5 %) of ATD patients and 30/41 (73.2 %) of CDG patients could be identified as ATD based on the quantitative information of the LC-MRM-MS test.

A



B

Cohort	Subtype	Antithrombin concentration Low	Normal	High
Control	-	-	35	2
ATD	All	76	11	4
	Type I	53	-	-
	Type II	23	11	4
	Subtype unknown	7	3	-
	HBS	4	4	-
	RS	2	3	4
	PE	10	1	-
CDG	All	30	9	2
	ALG12	1	-	-
	DPAGT1	1	-	-
	MPI	1	-	-
	PMM2	27	8	2
	SSR4	-	1	-

Figure 2. Quantitative analysis of AT in plasma of control, ATD and CDG samples by LC-MRM-MS. A) Concentrations of the analyzed samples, grouped per type. Dashed lines indicate reference intervals. Subtypes are specified for type II and CDG samples as indicated in the legend. **B)** Overview of the number of samples that are low, normal or high in concentration based on previously established reference intervals, grouped per type and subtype.

Molecular characterization

The qual/quant ratio, calculated by dividing the concentration of a qualifier peptide over the overall AT quantity, indicates how well an individual peptide concentration agrees with the overall AT concentration. A maximum deviation from the qual/quant ratio was established based on previously obtained precision data and empirical data (see also **Supplementary information** and [218]). Deviating qual/quant ratios are indicative of the presence of varying molecular proteoforms in a single sample. It was anticipated that deviating qual/quant ratios would only occur for type II ATD, as type I ATD mutations are believed to not be expressed and/or secreted. Indeed, deviating qual/quant ratios corresponding to the anticipated affected peptide (as based on genetic information) only occurred in 24 out of 38 (63.2 %) type II ATD samples; these patients likely express a variant AT proteoform (**Table 2**).

For these 24 samples the amount of wildtype proteoform that is present could be calculated based on the ratio of the total AT: % variant proteoform = (1 - qual/quant ratio) × 100 %, revealing proteoforms to be present at levels between 16 - 58 %. Of the 38 type II ATD samples, there were three mutation sites that each had three affected samples; namely Asn-225 (p.Asn224His), Glu-227 (p.Glu227Lys) and Arg-425 (p.Arg425del, p.Arg425His, p.Arg425Cys). These sites had mean qual/quant ratios ± SD of 1.35 ± 0.08, 0.72 ± 0.04 and 0.52 ± 0.07 respectively, indicating that (similar) mutations have a relatively consistent effect on the ratio of variant:wildtype proteoforms.

Of the 14 type II ATD samples that did not show deviating qual/quant ratios, nine samples harbored mutations outside of the assay coverage. Interestingly, 12 out of 14 unidentified samples were already flagged as being ATD based on their low concentration. The five samples which had mutations occurring in the monitored wildtype peptides also had low concentrations, likely caused by a low expression/secretion of the variant proteoform, hampering their identification by the qual/quant ratio.

Identification of specific proteoforms by variant peptides

To confirm the identity of the specific molecular variant, a confirmative approach was developed, in which the MS test was adjusted to measure variant peptides. In this pilot study, the genetic information available for each patient eased identification of the variant peptide. However, in future investigations the established variant peptide transitions may be directly applied without prior genetic knowledge or possible variants may be extracted from literature (as was previously applied in a case report [140]). Variant peptides could be identified in 22 ATD patients amounting to a total of 19 unique variant peptides (**Table 2**). The approach was applied to all samples, including those that were not indicated to have a qualitative defect based on the qual/quant ratio and type I ATD patients. Using this approach, ten patients that were not identified using the qual/quant ratio were found to carry a variant proteoform. When combined, the information from the qual/quant ratios and the variant peptide transitions identified a total of 34 patients as having a qualitative ATD using the LC-MRM-MS test.

Variant proteoforms were observed in four mutations classified as type I ATD (p.Phe155del, p.Lys157Arg, p.Phe271Ser, p.Leu441Pro), which were expected to merely exert a quantitative effect and no expression/secretion of the mutant proteoform. Similarly, variant peptides were observed in samples that did not show a deviating qual/quant ratio. This was the case for the p.Ala416Ser mutation, which was located outside of the coverage of the wildtype peptides, but for which the mutation created a new tryptic variant peptide that could be observed. Two mutations at site Met-283 (p.Met283Lys and p.Met283Val) should theoretically have led to a deviating qual/quant ratio for the wildtype peptide ADGES. Variant peptide could be observed for both samples which also presented with low overall AT concentrations of 0.57 and 0.72 µmol/L (for p.Met283Lys and p.Met283Val, respectively), strengthening our initial hypothesis that the low overall AT concentration in these samples is caused by a low concentration of variant proteoform hampering the identification of the mutation through qual/quant ratios.

Table 2. Identification of qualitative deficiencies and variant proteoforms.

patient	mutation ◆	deficiency subtype	qual/quant ratio or glyco/quant ratio (Peptide) ▲	variant peptide detected ▲
1	p.Val30Glu	unknown*	0.49 (HGSPV)	Yes
2	p.Val30Glu p.Arg425Cys	unknown* type II RS	0.69 (HGSPV) 0.58 (SLNPN)	Yes
3	p.Arg45Trp	type II HBS	0.49 (HGSPV) 0.47 (DIPMN)	No
4	p.Pro73Leu	type II HBS	0.42 (IPEAT)	Yes
5	p.Arg79Cys	type II HBS	1.27 (IPEAT)	Yes
6	p.Arg79His	type II HBS	1.32 (IPEAT)	Yes
7	p.Leu131Phe	type II HBS	0.67 (GP-LGACN)	Yes
8	p.Val137Ala	type II PE	0.81 (GP-LGACN)	Yes
9	p.Phe155del	type I	-	Yes
10	p.Lys157Arg	type I	-	Yes
11	p.Arg177Cys	type II HBS	0.48 (LVSAN)	Yes
12	p.Gly199Arg	type II PE	0.98 (GP-SLTFN) 0.79 (LQPLD)	No
13	p.Asn224His	type II	0.71 (GP-WVSNK)	No
14	p.Asn224His	type II	0.69 (GP-WVSNK)	No
15	p.Asn224His	type II	0.77 (GP-WVSNK)	No
16	p.Glu227Lys	type II	1.28 (GP-WVSNK)	No
17	p.Glu227Lys	type II	1.44 (GP-WVSNK)	No
18	p.Glu227Lys	type II	1.32 (GP-WVSNK)	No
19	p.Asn240Lys	type II RS	-	Yes
20	p.Phe271Ser	type I	-	Yes
21	p.Lys273Glu	type II PE	0.79 (ELFYK)	No
22	p.Met283Val	type II RS/PE	-	Yes
23	p.Met283Ile	type II RS/PE	0.83 (ADGES)	Yes
24	p.Met283Lys	type II PE	-	Yes
25	IVS2-2 A>T p.Ala416Ser	type I (type II RS)*[2]	- -	- Yes
26	p.Ala416Ser	type II RS	-	Yes
27	p.Ala416Ser p.Ser397Leu	(type II RS)*[2] type I	- -	Yes -
28	p.Arg425del	type II RS	0.55 (SLNPN)	No
29	p.Arg425His	type II RS	0.58 (SLNPN)	No
30	p.Arg425Cys	type II RS	0.44 (SLNPN)	No
31	p.Ser426Leu	type II RS	0.43 (SLNPN)	Yes
32	p.Pro439Thr	type II PE	0.79 (ANRPF)	Yes
33	p.Leu441Pro	type I	-	Yes
34	p.Pro461Ser	type II PE	0.84 (VANPC)	Yes

◆ Mutation as identified by genetic analysis. ; ▲ as detected by LC-MRM-MS analysis; * note that the p.Val30Glu mutation is expected to cause a transient quantitative type I deficiency, as described by Navarro-Fernandez et al. (2016) [265]. However, presence of variant proteoform suggests a qualitative type II deficiency. Therefore, subtype is marked as unknown. *[2] These samples contained double mutations of type I and type IIRS, but are classified as type I due to the dominant effect of the type I mutation.

Identification of abnormal glycosylation

Beyond *SERPINC1* mutations, ATD may be caused by abnormal glycosylation of AT, for instance as a result of CDGs. Whereas mutations are fixed variations, glycosylation is a variable process which may be influenced by CDG subtypes or temporal fluctuations. Thus, the impact of CDG on the glycopeptides was expected to show larger variability than that of mutations on regular peptides. To characterize the degree of hypoglycosylation of AT proteoforms in CDG patient plasma the four glycosylation sites of AT were monitored by measuring four glycopeptides, each targeting a single glycosylation site. One glycopeptide, GP-KANK, monitors the site (Asn-167) known to be hypoglycosylated naturally [22], leading to the distinction between two main proteoforms of AT; α-AT and β-AT, the latter contributing to ~10 % of the total AT in healthy adults [190]. Given the natural variability in this glycosylation site, it was anticipated that CDGs would have the largest effect on GP-KANK. Furthermore, a previous study showed that a second glycosylation site may be affected in CDG patients [60].

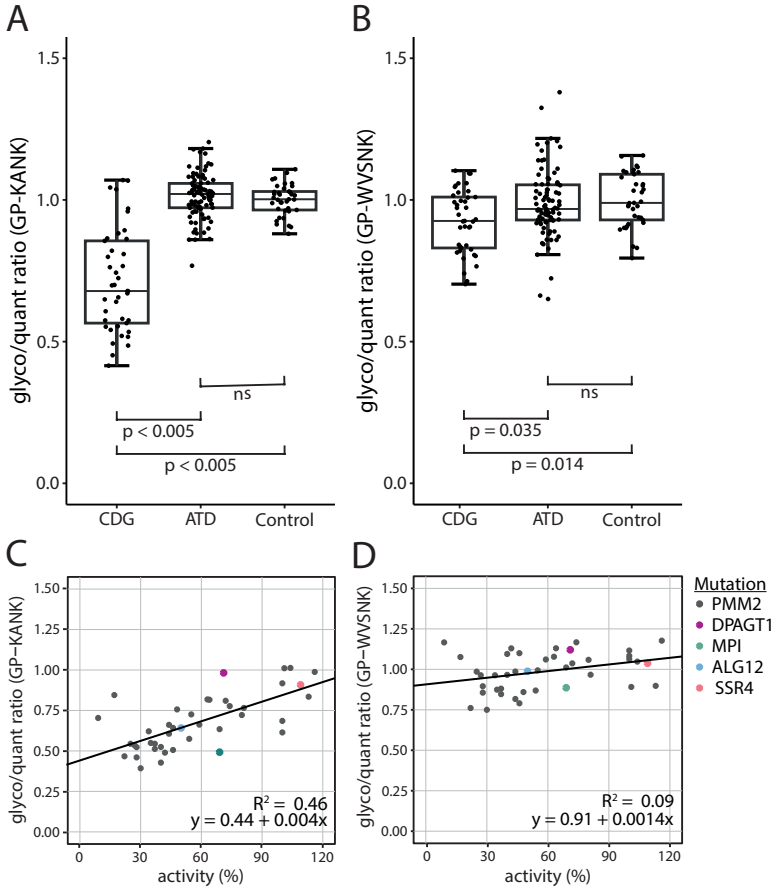


Figure 3. Analysis of glyco/quant ratio data for GP-KANK and GP-WVSNK. A-B) Boxplots for the normalized glyco/quant ratios for GP-KANK and GP-WVSNK, respectively. ns = not significant. **C-D)** Linear regression analysis for glyco/quant ratios versus the activity for GP-KANK **(C)** and GP-WVSNK **(D)**, respectively.

Significant differences were observed in the glycosylation of CDG patients versus controls for two glycopeptides (**Figure 3A-B**). Indeed, the largest difference was observed for glycopeptide KANK (GP-KANK), with a median normalized qual/quant ratio of 0.68 in CDG patients compared to 1 in controls ($p < 0.005$). In agreement with the study by de la Morena-Barrio et al. [60] a second significant

change was found for GP-WVSNK, reflecting Asn-224, which showed a median normalized qual/quant ratio of 0.93 in CDG patients compared to 0.99 in controls (p = 0.014). Interestingly, 27 out of 41 (65.9 %) CDG patients had a GP-KANK ratio below the lowest ratio found in healthy controls (0.83). As expected, a larger variation in the glycosylation was observed as compared to mutations, with the qual/quant ratio for GP-KANK ranging between 0.42 – 1.07 in CDG patients. Samples with an aberrant glycopeptide/quant ratio were marked to likely harbor a glycosylation defect.

Correlation between glycosylation and activity

Incorrect glycosylation of AT affects its functionality [60]. To investigate this, the correlation between the functionality of AT based on the diagnostic activity test was compared to the glycopeptide/quant ratios for CDG patients. Significant linear correlations with activity were found for both GP-KANK ($p < 0.005$) and GP-WVSNK ($p = 0.03$) ratios when assessed individually (**Figure 3C-D**), supporting a correlation between the degree of wildtype glycosylation and AT functionality. However, hypoglycosylation also affects the overall AT concentration due to lowered secretion. Thus, linear regression analysis of the overall AT concentration, GP-KANK concentration, and GP-KANK ratio versus the activity were performed, revealing GP-KANK to be the best predictor of activity (as reflected by higher r-squared values of 0.51 and 0.46 for GP-KANK concentration and ratio versus 0.42 for overall AT concentration). Combining these variables in a multiple linear regression mode revealed that the best model (based on the lowest Akaike Information Criterion) for predicting the activity was: Activity = -18.57 + 26.58 * quantifier + 78.32 * GP-KANK (see also **Supplementary Table 1**). Together, this indicates that the LC-MRM-MS results may be used to (partially) predict the functionality of AT.

Identification of ATD patients by LC-MRM-MS versus activity test

In a diagnostic setting, the AT activity test serves as a first line test to identify ATD in patients suspected of having a thrombophilia. A cut-off of 80 % activity is often applied (equal to the RI of the here applied activity test), classifying samples below 80 % as ATD. The diagnostic sensitivity of the activity tests and the LC-MRM-MS test for the ATD samples studied here are shown in **Table 3**. Interestingly, both tests performed well for type I deficiencies, with 98.1 % and 100 % identified samples for the activity and LC-MRM-MS test, respectively. However, the activity test only identified 56.8 % of the

Table 3 Diagnostic sensitivity of the activity test and the LC-MRM-MS test for ATD patients.

	Activity		LC-MRM-MS					
			Quantitative		Qualitative		Combined	
	N	%	N	%	N	%	N	%
ATD	73/90*	81.1	76/91	83.5	34/91	37.4	91/91	100
Type I	52/53	98.1	53/53	100	6/53*[2]	11.3	43/43	100
Type II	21/37*	56.8	23/38	60.5	28/38	73.7	39/39	100
- HBS	6/8	75.0	4/8	50.0	6/8	75.0	8/8	100
- PE	4/11	36.4	10/11	90.9	7/11	63.6	11/11	100
- RS	4/8*	50.0	2/9	22.2	8/9	88.9	9/9	100
- NA	7/10	70.0	7/10	70.0	7/10	70.0	11/11	100

* Missing activity value for one sample. *[2] Two samples contained a combination of a type I and type II mutation and were classified as type I due to the dominant phenotype of this mutation.
Identification by the activity test entailed an activity <80 %. For LC-MRM-MS identification, a quantitative defect entailed a concentration below the reference intervals (1.07 µmol/L). A qualitative defect entailed a deviating qual/quant ratio and/or a measured variant peptide.

type II deficiencies, with the lowest rate found for type II PE deficiencies (36.4 %), which is in line with previous studies reporting low sensitivity of activity tests for type II (HBS) ATD [14, 202]. In stark contrast, the LC-MRM-MS test identified 60.5 % of the type II deficiencies based on AT quantity alone, rising to 100 % when combining the quantitative and qualitative information. Of note, eight of the 17 patients (47.0 %) missed by the activity test did suffer from a thrombotic event, indicating the clinical relevance of identifying these patients.

For CDG patients, screening for ATD (among other thrombophilia's) is suggested to be performed at regular intervals or under provoking conditions [263, 264]. It is not known nor expected that CDG patients are at constant risk for thrombophilia, as their coagulation parameters may fluctuate over time, depending on the extent to which the glycosylation deficiency may be remitted. Therefore, the LC-MRM-MS test was compared to the activity test (see also **Supplementary Table 2**) for identifying those at risk. The LC-MRM-MS test identified a slightly larger percentage of patients as ATD compared to the activity test (87.8 % versus 75.6 % for the LC-MRM-MS and activity test, respectively. Interestingly, the five patients not identified by the MS test also had normal values (ranging between 100 and 116 %) in the activity test and only slightly increased aberrant transferrin values (CDG-marker, see **Supplementary Table 3**), suggesting that these patients were at low risk for thrombosis during sampling.

Overall, in this cohort of ATD and CDG patients, the LC-MRM-MS test provides high diagnostic sensitivity, with increased sensitivity in the case of ATD patients compared to the activity test, and comparable sensitivity compared to contemporary tests for identifying discrepancies in CDG patients.

Discussion

Diagnosing ATD by mass spectrometry-based precision diagnostics is a valuable concept, but evidence for the clinical performance of this test was needed. There is a need to diagnose ATD more accurately in patients with ATD caused by a *SERPINC1* defects, specifically in the ATD-type II HBS subtype. Moreover, availability of a molecular diagnosis could enable studies towards targeted treatments. To address this need, AT was evaluated using our LC-MRM-MS test in a cohort of 91 patients with ATD caused by *SERPINC1* defects. By combining quantitative and qualitative data, the tests' diagnostic sensitivity for ATD outperformed that of the current activity test with an identification rate of 100 % versus 81.1 % for the activity test (**Table 3**). The test could also identify type II HBS mutations known to be difficult to diagnose through activity tests, such as p.Pro73Leu (AT Basel), p.Leu131Phe (AT Budapest III) and p.Arg79His (AT Padua). Compared to the activity test, the LC-MRM-MS test identified additional patients and enabled ATD subtyping, highlighting its potential to improve the current diagnostic pathway. Taken together, this study strongly supports the notion that molecular characterization of AT proteoforms improves diagnostic performance and enables identification of patients with *SERPINC1* defects in the diagnostic "grey zone" which are at risk for misdiagnosis.

AT carries 4 N-glycosylation sites, and altered glycosylation has been reported to influence AT activity. Consequently, glycosylation was included in our LC-MRM-MS test. Patients with CDG carry altered glycosylation and can have constitutive or transient ATD. Current ATD activity tests do not target AT glycosylation specifically, whilst studies investigating ATD in CDG patients found low AT activity levels in approximately 80 % of PMM2-CDG patients leading to high occurrence of VTE (~10 %) [263, 266]. To assess the performance of our LC-MRM-MS test in the identification of ATD in CDG patients, AT was measured in a cohort of 41 patients with CDG. We identified a significant difference in the glycopeptide/quant ratios of patients compared to controls for GP-KANK and GP-WVSNK (representing glycosylation sites Asn-167 and Asn-224). Regarding the diagnostic sensitivity, the test classified a slightly higher percentage of patients in the ATD group compared to the activity test (87.8 % versus 75.6 %) indicating that the test performs similar to the activity test with regards to identifying CDG patients at risk for thrombosis. As we currently do not have a "gold standard" for CDG patients,

further evaluation in relation to VTE incidence would be warranted in the CDG population.

Beyond diagnostic performance, our findings open the discussion on the definition of type I and type II ATD. Classically, type I AD is characterized by low activity and low antigen levels, and type II ATD by low activity and normal to slightly lower antigen levels. In this study 23 out of 38 (60.5 %) type II ATD patients showed low quantities of AT, with many of the samples even showing similar concentrations to type I patients. Because type I ATD is typically regarded as more pathogenic than type II, these low quantities suggest that certain type II deficiencies should likely be regarded equally pathogenic as type I ATD. Conversely, variant peptides were detected in ATDs currently classified as type I, specifically p.Phe155del, p.Lys157Arg, p.Phe271Ser and p.Leu441Pro. These results suggest that certain molecular variants must be reclassified and that the current classification system only provides a crude classification, likely not in line with the actual disease severity. . Importantly, the clinical phenotype and severity of specific type II ATD at the molecular level has not systematically been studied. However, the presence of variants could lead to either ameliorating effects, such as retained activity in the β-AT proteoform of certain HBS mutations [45], or dominant negative effects [260, 267]. Thus, proteoform analysis allows establishment of a more refined classification system that may better resolve patients at high and low risk for thrombosis based on a molar quantitation combined with proteoform identification.

There is clear potential for this test to improve our understanding of ATD, but for it to provide added value in a diagnostic setting it must improve a current clinical gap. For ATD this entails reduction of diagnostic uncertainty especially for patients with borderline activity values and/or improved risk stratification of patients. For both instances, the test could be applied as an "add-on" test to the activity test to 1) verify the absence/presence of ATD in borderline cases, and 2) provide molecular information to better estimate thrombosis risk. Although the latter objective requires longitudinal studies investigating the thrombosis risk of specific AT variants, the feasibility of the first objective was partially verified in this study. Part (47.0 %) of the additional ATD patients identified by LC-MRM-MS, but not by the activity test, already suffered from a clinical event, highlighting the inadequacy of the current test. Furthermore, this is a clear example of conventional "imprecision diagnostics" in our reactive healthcare system, in which patients first have to undergo a clinical event, potentially causing lifelong damage, for a disease to be uncovered [1, 268]. Current guidelines are hesitant towards thrombophilia screening, as current thrombophilia tests only identify a cause in less than half of the patients and knowledge on the cause often does not alter the treatment strategy [149]. Thus, with the AT LC-MRM-MS test providing both a more refined approach to identifying ATD and offering molecular insight into the disease, driving therapeutic developments, the thrombophilia community should be aware that an ounce of prevention by next-generation test screening can be worth a pound of cure.

The results warrant exploration of the full clinical performance in a longitudinal setup, for instance by comparing current practice (activity test) with the comparator test (LC-MRM-MS) in patients presenting with thrombophilia. As was the case for the current study, the low prevalence of ATD in the general population (1:400 to 1:600) may complicate the design of such a diagnostic accuracy study potentially requiring alternative approaches [18, 269]. However, provided sufficient ATD cases are enrolled, and patient follow-up is sufficiently long, a well-designed diagnostic accuracy study may also allow for the development of risk stratification based on ATD molecular variants.

Although we believe that protein-level diagnostics offers valuable information that cannot be provided by functional or genetic testing, the LC-MRM-MS test is not omnipotent. The test performs well in identifying quantitative ATD but cannot discriminate between hereditary type I ATD and transient ATD. Consequently, exclusion of transient factors, presence of family history and/or genetic sequencing is still required to definitively classify a hereditary type I ATD, although these are also required when

using activity tests. Furthermore, analogous to clinical care pathways, which must be improved continuously with novel insights and technologies, the LC-MRM-MS test also requires continuous development. In this study, the genetic and clinical information of the samples was essential to establish and verify a data analysis strategy and develop variant proteoform specific transitions. The test now identifies many ATD mutations down to the amino acid levels, such as AT Basel, AT Padua I and AT Budapest III, which occur in high frequencies in Europe [89, 90]. Additionally, variant peptides for type II ATD mutations that were not present in the current cohort can be included for identification upon encounter. Of note, as the LC-MRM-MS technique is only established in in specialized laboratories, centralized expert centers are envisioned to facilitate access to the in-house developed molecular AT-test, which only requires a small sample volume (< 30 uL), provides stable results over multiple freeze/thaw cycles and costs currently < $100 per test .

With the future vision on P5 medicine often focused on novel treatment strategies, such as gene therapy [270], we must not overlook that the clinical pathway is only as strong as each of its elements. For ATD, the current clinical pathway relies on a diagnostic strategy that may misdiagnose patients and creates a clinically heterogenous patient stratification. The LC-MRM-MS test refines the ATD diagnosis, enabling an improved diagnostic pathway and holds potential for improved patient stratification. To accomplish better patient management, tailored interventions targeting specific patient groups must be developed, for which molecular insight is essential, analogous to e.g. the oncology field where genotyping cancers allows personalized treatment [271]. Taken together, the LC-MRM-MS test enables detection and quantification of harmful AT-proteoforms which diminishes diagnostic uncertainty that hemostasis doctors face in the current ATD clinical care pathway, thereby bringing ATD Precision Diagnostic and Healthcare into the 21st century.

Acknowledgments

We thank all patients who contributed to this study. Furthermore, we thank ThermoFisher for providing the CaptureSelect α-Antithrombin used in the assay. This project received funding from the European Union's Horizon 2020 Research and Innovation program under the Marie Skłodowska-Curie grant agreement (843615), Fundación Séneca (21886/PI/22) and the Instituto de Salud Carlos III (ISCIII) through project PI21/00174 co-funded by the European Union. MEM-B has a Ramon y Cajal contract (RYC2021-031000-I, Ministerio de Ciencia, Spain).

Supplementary Table 1. Multiple linear regression models for explaining AT activity by MS results.

	Model 1			Model 2			Model 3			Model 4		
Predictors	Estimates	CI	p	Estimates	CI	p	Estimates	CI	p	Estimates	CI	p
(Intercept)	-6.98	-65.28 – 51.32	0.810	-20.88	-79.48 – 37.71	0.475	-18.57	-45.68 – 8.54	0.174	10.62	-67.87 – 89.11	0.785
Quantifier	28.06	-0.70 – 56.83	0.056				26.58	-1.11 – 54.27	0.059	-6.81	-95.50 – 81.88	0.877
GP-KANK	81.30	23.18 – 139.42	0.007	117.14	70.45 – 163.83	<0.001	78.32	22.37 – 134.28	0.007	35.68	-85.68 – 157.04	0.555
GP-WVSNK	-15.13	-82.30 – 52.03	0.651	-0.16	-67.95 – 67.62	0.996						
Quantifier * GP-KANK										45.65	-69.47 – 160.76	0.427
Observations	41			41			41			41		
R^2 / R^2 adjusted	0.523 / 0.484			0.473 / 0.445			0.520 / 0.495			0.528 / 0.490		
AIC	371.499			373.615			369.729			371.020		

CI = confidence interval, p = p-value, AIC = Akaike information criterion

Data analysis process

In short, the molecular characterization of AT by LC-MRM-MS is based on three layers of information and built from data on the (molar) concentration of the 23 AT peptides monitored in the test, as described previously, with one adjustment concerning the maximal allowed deviation (Kruijt et al., manuscript submitted). First, the overall concentration of AT is verified by comparing the concentrations of three quantitative peptides and reporting a single verified peptide concentration that is representative for the total AT proteoform mixture. Second, all remaining peptide concentrations are compared to the total concentration, giving rise to the qual/quant or glycopeptide/quant ratio. Major deviations from 1 are indicative of mutations or changes in the glycosylation. Deviation limits were based on the within laboratory precision of each peptide observed during the analytical validation [218], and calculated as following:

$$maximal\ allowed\ deviation = 1 \pm \frac{(precision\ quantifying\ peptide + precision\ qualifying\ peptide) \times 1.2}{100}$$

The lower limit was applied to all peptides and the upper limit was only applied to peptides IPEAT and GP-WVSNK, as these were the only observed peptides for which a mutation could lead to increased concentrations of wildtype peptide. For GP-KANK the deviation limit (0.85) was based on Youden's index due to the natural variability in the α- and β-proteoforms of AT affecting the glycopeptide/quant ratio. These limits did result in false positives for specific peptides in the current data set, likely caused by the storage condition affecting peptides sensitive for oxidation. MS-transitions that monitored specific variant peptides were developed and applied to 1) verify aberrant qual/quant ratio results (thereby excluding false positives); and 2) identify molecular proteoforms down to the amino acid level.

Supplementary Table 2. Diagnostic sensitivity of the activity test and the LC-MRM-MS test for CDG patients.

	Activity		LC-MRM-MS					
			Quantitative		Qualitative		Combined	
	N	%	N	%	N	%	N	%
CDG	31/41	75.6	30/41	73.2	35/41	85.4	36/41	87.8
- PMM2	28/37	75.7	27/37	73.0	33/37	89.2	33/37	89.2
- ALG12	1/1	100	1/1	100	1/1	100	1/1	100
- MPI	1/1	100	1/1	100	1/1	100	1/1	100
- DPAGT1	1/1	100	1/1	100	0/1	0	1/1	100
- SSR4	0/1	0	0/1	0	0/1	0	0/1	0
- NA	7/10	70.0	7/10	70.0	7/10	70.0	11/11	100

For LC-MRM-MS identification, a quantitative defect entailed a concentration below the reference intervals. A qualitative defect entailed a deviating glycopeptide/quant ratio and/or a measured variant peptide.

Supplementary Table 3. Transferrin values for CDG-patients not identified by MS.

Identified by MS	Ranges asialo transferrin values (%)	Ranges of ratio disialo/trisialo transferrin values
Yes	0.07 – 28.42	0.11 – 62.17
No	0.00 – 2.75	0.30 – 2.59

Reference values for asialo % and disialo/trisialo ratio are 0.5+-0.3 % and 0.40 +-0.09, with values above 1 considered pathogenic for both values.