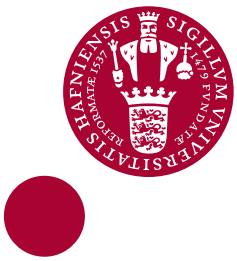

Does diabetes modify the effect of heparin on plasma proteins? - A proteomic search for plasma protein biomarkers for diabetes-related endothelial dysfunction

Soerensen, Mette; Debrabant, Birgit; Halekoh, Ulrich; Møller, Jacob Eifer; Hassager, Christian; Frydland, Martin; Hjelmborg, Jacob; Beck, Hans Christian; Rasmussen, Lars Melholt

Published in:
Journal of Diabetes and its Complications

DOI:
10.1016/j.jdiacomp.2021.107906

Publication date:
2021

Document version
Publisher's PDF, also known as Version of record

Journal of Diabetes and Its Complications 35 (2021) 107906

Contents lists available at ScienceDirect

Journal of Diabetes and Its Complications

journal homepage: WWW.JDCJOURNAL.COM

Does diabetes modify the effect of heparin on plasma proteins? - A proteomic search for plasma protein biomarkers for diabetes-related endothelial dysfunction

Mette Soerensen [a,b,c,*,1], Birgit Debrabant [a,1], Ulrich Halekoh [a], Jacob Eifer Møller [d,e,f], Christian Hassager [e,f], Martin Frydland [e,f], Jacob Hjelmborg [a], Hans Christian Beck [b], Lars Melholt Rasmussen [b]

[a] Epidemiology, Biostatistics and Biodemography, Department of Public Health, University of Southern Denmark, J.B. Winsløws Vej 9B, 5000 Odense C, Denmark
[b] Center for Individualized Medicine in Arterial Diseases, Department of Clinical Biochemistry and Pharmacology, Odense University Hospital, J.B. Winsløws Vej 4, 5000 Odense C, Denmark
[c] Department of Clinical Genetics, Odense University Hospital, J.B. Winsløws Vej 4, 5000 Odense C, Denmark
[d] Department of Clinical Cardiology, Odense University Hospital, J.B. Winsløws Vej 4, 5000 Odense C, Denmark
[e] Department of Cardiology, Rigshospitalet, Blegdamsvej 9, 2100 Copenhagen Ø, Denmark
[f] Department of Clinical Medicine, University of Copenhagen, Blegdamsvej 3B, 2200 Copenhagen N, Denmark

ARTICLE INFO

Article history:
Received 11 November 2020
Received in revised form 11 February 2021
Accepted 7 March 2021
Available online 17 March 2021

Keywords:
Diabetes
Heparin
Proteome data
Mediation
Interaction models
Endothelial dysfunction

ABSTRACT

Aim: Heparin administration affects the concentrations of many plasma proteins through their displacement from the endothelial glycocalyx. A differentiated protein response in diabetes will therefore, at least partly, reflect glycocalyx changes. This study aims at identifying biomarkers of endothelial dysfunction in diabetes by statistical exploration of plasma proteome data for interactions between diabetes status and heparin treatment.
Methods: Diabetes-by-heparin interactions in relation to protein levels were inspected by regression modelling in plasma proteome data from 497 patients admitted for acute angiography. Analyses were conducted separately for all 273 proteins and as set-based analyses of 44 heparin-relevant proteins identified by gene ontology analysis and 42 heparin-influenced proteins previously reported.
Results: Seventy-five patients had diabetes and 361 received heparin before hospitalization. The proteome-wide analysis displayed no proteins with diabetes-heparin interaction to pass correction for multiple testing. The overall set-based analyses revealed significant association for both protein sets (p-values$<2^{*}10^{-4}$), while constraining on opposite directions of effect in diabetics and none-diabetics was insignificant (p-values = 0.11 and 0.17).
Conclusions: Our plasma proteome-wide interaction approach supports that diabetes influences heparin effects on protein levels, however the direction of effects and individual proteins could not be definitively pinpointed, likely reflecting a complex protein-basis for glycocalyx dysfunction in diabetes.

1. Introduction

The vascular endothelium is an important multifunctional tissue and is essential for the modulation of vascular function and structure and consequently the general homeostasis of the body.[1] Lining the luminal side of the vascular endothelium, the endothelial glycocalyx serves as a permeability barrier shielding the vascular wall from direct exposure to the blood flow, yet it is also important for signalling between the endothelial cells and the intravascular lumen.[2] The glycocalyx is constituted of a mixture of mostly glycoproteins, glycolipids and highly negatively charged membrane-bound proteoglycans, consisting of glycosaminoglycan, for example heparan sulphate. In case the structure of the glycocalyx is disrupted, it might lead to an increased permeability and consequently to exposure of the endothelial cells, e.g. increased inflammation[3] and subsequently vascular complications to diseases such as hypertension, chronic kidney disease and diabetes.[4–6]

Of specific relevance to diabetes, Deckert et al. thirty-one years ago formulated the Steno hypothesis[7] stating that the albuminuria and vascular dysfunction observed in diabetes patients could be due to a

Abbreviations: FDR, false discovery rate; GO, Gene Ontology.
Declaration of competing interest: None.

* Corresponding author at: Epidemiology, Biostatistics and Biodemography, Department of Public Health, University of Southern Denmark, J.B. Winsløws Vej 9B, 5000 Odense C, Denmark.
E-mail addresses: msoerensen@health.sdu.dk (M. Soerensen), bdebrabant@health.sdu.dk (B. Debrabant), uhalekoh@health.sdu.dk (U. Halekoh), jacob.moeller1@rsyd.dk (J.E. Møller), hassager@dadlnet.dk (C. Hassager), martin.steen.frydland.01@regionh.dk (M. Frydland), JHjelmborg@health.sdu.dk (J. Hjelmborg), hans.christian.beck@rsyd.dk (H.C. Beck), lars.melholt.rasmussen@rsyd.dk (L.M. Rasmussen).
[1] These authors contributed equally to the work.

https://doi.org/10.1016/j.jdiacomp.2021.107906

generalised reduction of negative charges of the extracellular matrix and plasma membranes, as a result of the loss of heparan sulphate (an important component of the glycocalyx), leading to changes in permeability of the vascular endothelium. Since then, molecular aspects of endothelial dysfunction in diabetes patients have been studied intensively, including the endothelial glycocalyx and its connection to aspects concerning oxidative stress, the role of nitric oxide and changes in inflammatory and vasomotoric processes.[8] Regarding glycocalyx disintegration, diabetes patients have been reported to have lower glycocalyx volume as compared to controls[4] and the disruption of the glycocalyx and factors potentially maintaining its stability and repair have been suggested as an important aspect of the cardiovascular complications in diabetes.[6] Good clinical markers for the putative effect of diabetes on glycocalyx and endothelial cell function are, however, unfortunately largely absent.

Several studies have shown that heparin injection (a widely used anticoagulant for treatment of among others thrombosis) acutely affects the concentrations of many of the proteins measured in blood samples,[9,10] as heparin competes with the binding of these proteins to proteoglycans in the endothelial glycocalyx, washing them into the circulation.[11] The plasma concentration of such proteins after heparin administration does therefore partly reflect the state of the glycocalyx and will change in situations with altered glycocalyx-binding capacity. As stated above, glycocalyx changes have been suggested, as part of the general endothelial dysfunction in diabetes[8] and we therefore hypothesize that the presence of diabetes interacts with the effect of heparin on plasma proteins, and that proteins which display significance for such interaction term (between diabetes and heparin treatment) may be putative biomarkers for endothelial dysfunction.

Consequently, the aim of the present study was to investigate whether diabetes status modifies the heparin induced protein shedding from the glycocalyx. For this aim, proteome data derived from plasma samples taken from 497 patients admitted for acute angiography,[12] of which some were treated with heparin before hospitalization and some had known diabetes, were investigated; the modifying effect of diabetes on the association between heparin treatment and protein levels was investigated by protein-wise regression analysis including an interaction term between heparin and diabetes. Furthermore, set-based analyses were applied to proteins with known relation to the glycocalyx to search for consistent interaction within the combined sets. Especially, we considered proteins identified by Gene Ontology (GO) terms as related to heparin and proteins for which the level has previously been found to change after heparin treatment in longitudinal studies.[10] Finally, we applied model selection to predict the presence of heparin based on multiple proteins and investigated whether protein-diabetes interactions contributed to the prediction.

Hence, this study aimed at identifying proteins related to the glycocalyx of specific relevance to diabetes patients holding heart complications and hence point to markers for endothelial injury in such patients.

2. Subjects, materials and methods

2.1. The study population and proteome data

The study population included 550 participants randomly selected from the Prediction and Risk assEsment in patients aDmitted to aCute coronary angiography and developmenT of Cardiogenic Shock (PREDICT-CS) study[12] admitted to Odense University Hospital for urgent coronary angiography with suspected ST-elevation myocardial infarction. From ambulance and patient files it was possible to obtain the clinical characteristics relevant for the present study for 497 patients; 361 (72.6%) had received heparin treatment before admission to the hospital and 75 (15.1%) patients were known to have diabetes. Of these, 4 patients (5.3%) had diabetes type 1, 69 patients (92%) had diabetes type 2 and 2 patients (2.7%) were registered as unknown type of diabetes. The average blood glucose level of the study population was 9.58 mmol/l (standard deviation (SD) = 4.53, $N = 399$), while it was 8.91 mmol/l (SD = 3.87, $N = 332$) for the non-diabetics and 12.9 mmol/l (SD = 5.96, $N = 67$) for the diabetics. The clinical characteristics included in the present study can be seen in Table 1, while additional available clinical information for the study population can be found in the Supplementary Table 1 of the Supplementary Material. Just prior to the coronary angiography procedure, a standard blood sample and an EDTA blood sample were collected and from the latter plasma were obtained after centrifugation. The plasma samples were used for proteomic analysis: all details related to sample preparation, labelling with mass tags, fractionation of tagged peptides, nano LC-MS/MS analysis and protein identification are described in Beck et al. 2018[10] and the proteome data can be found via ProteomeXchange with identifier PXD008468[dataset, 13]. Protein measurements are given as ratios of the actual amount of protein in a plasma sample to the average amount in a calibrator-plasma-pool generated from patients, who did not receive heparin, as described in detail by Beck et al. 2018.[10] The study was approved by the local ethics committee for Region Hovedstaden (Capital Region of Denmark) (approval H-2-2014-110) and the study was conducted in accordance with the Declaration of Helsinki. Informed consent was obtained from all patients or their next of kin, as well as the patient's general practitioner, in agreement with Danish legislation.

2.2. Preparation and imputation of data

Initially measurements were available for 723 proteins. Outlying values, defined as protein measurements >80, were set to be missing values and proteins with a call rate < 50% were excluded, leaving 273 proteins for analysis. The data for these 273 proteins were log-transformed and missing values were imputed based on observations from other proteins. The number of imputations was set to five. Finally, measurements were standardized for each protein using the observed standard deviation and mean (average over the different imputations). Correspondingly, multiple imputation was conducted to produce five imputations for missing clinical data.

2.3. Identification of proteomic feature sets for set-based analyses

With the aim to investigate proteins of specific relevance for heparin binding and hence shredding from the glycocalyx, we performed two types of set-based analyses of subsets of proteins, i.e. proteins identified as relevant via search for Gene Ontology (GO) terms or proteins associated with heparin in an earlier study of 9 individuals, who had been followed longitudinally after heparin treatment.[10]

2.3.1. Proteins identified via Gene Ontology terms related to heparin or heparan binding

A search in the GO database[14] and the QuickGo database[15] with the search term 'heparin' revealed three GO terms related to either heparin binding (GO:0008201), heparan sulfate binding (GO:1904399) or heparan sulfate proteoglycan binding (GO:0043395). The three GO terms contained 193, 5 and 21 protein annotations for *Homo sapiens*,

Table 1
Clinical characteristics of the study population.

	Study population
No. individuals	497
Males (%)	354 (71.2)
Mean age (SD)	64.44 (13.96)
Mean body mass index (SD)	26.79 (4.63)
Diabetes (%)	75 (15.1)
Heparin treatment before hospitalization (%)	361 (72.6)

Note: No.: number of, SD: standard deviation.

respectively. These annotations were verified using the STRING database[16] and corresponded to 182 unique human proteins. The overlap with the 273 proteins of the proteome dataset of the present study was investigated, leading to 44 proteins (see Supplementary Table 2), which were used for the sub-analysis in the set-based analyses (see below).

2.3.2. Proteins previously shown to be associated with heparin in a longitudinal study

We selected proteins that have been shown to be associated with heparin in an earlier study of 9 patients admitted to acute coronary angiography.[10] Especially, the study assessed the effect of heparin 2, 15 and 60 min after heparin administration and compared protein levels after administration with baseline levels for each time point and each protein.[10] Based on *p*-values from this previous study, for our article we calculated false discovery rates (FDR)[17] considering all three time points and all proteins jointly. We identified those 64 proteins that reached an FDR ≤ 0.05. Only 42 proteins overlapped with the proteins of this study and formed the basis for the set-based analysis (see Supplementary Table 3).

2.4. Statistical analyses

We studied the association of protein levels with diabetes and heparin by univariate, as well as set-based analyses. Our main interest focused on the interaction between diabetes and heparin relative to proteins, that is, we hypothesized that presence of diabetes changes the effect of heparin (or vice versa that heparin changes the impact of diabetes) on protein levels. Finally, we embedded our research question into the context of predicting heparin administration using multi-proteomics-interaction-modelling and investigated whether diabetes changes the way in which proteins predict heparin. Diabetes and heparin were used as dichotomous variables corresponding to the presence of diabetes (yes/no) and the administration of heparin (yes/no).

2.4.1. Univariate regression modelling

For each protein separately, we regressed protein levels on diabetes and heparin including an interaction term in the model. Further covariates were age, sex and body mass index (BMI). Effect estimates from the five imputations are combined according to Rubin's formula.[18] Reported *p*-values correspond to the null hypothesis of no interaction. Reported coefficients correspond to initially log-transformed and standardized protein levels (cp. Section 2.2. Preparation and imputation of data). The log-transformation is intended to improve normality assumptions underlying regression-based testing. Correction for multiple testing was conducted by the Benjamini-Hochberg method[17] obtaining FDR corrected *p* values. A cut-off of 0.05 was used in order to declare a protein to be a finding. For selected proteins, marginal effects, i.e. the expected difference in outcome between diabetic and non-diabetic patients while holding the other variables unchanged, were plotted depending on heparin administration.

2.4.2. Set-based analyses

Next, we tested the interaction hypothesis considering sets of proteomic features. This comprises two advantages. Firstly, it circumvents the multiple testing problem when testing single feature sets in a confirmatory way instead of hundreds of individual proteins. Moreover, when several proteins are considered jointly, individual minor effect sizes could jointly create evidence against the null hypothesis H_0 of no interaction in any of the proteins of the set. Apart from our focus on interaction, similar set-based tests are known from gene set analysis, which is widely applied to especially genome-wide association studies and genome-wide gene-expression data.[19,20]

We considered two alternative hypotheses: the general alternative $H_{a;general}$ stating that an interaction is present for at least one of the proteins of the set, and the more directed alternative $H_{a;directed}$ stating that there is an interaction for some of the proteins and it is such that diabetes weakens or reverses the effect of heparin on the proteins of the sets. The general hypothesis assumes no specific directions of effect for the proteins, while the directed hypothesis addresses proteins for which the effect of heparin on protein levels differs between diabetic and non-diabetic individuals in a way that the effect of heparin is weakened or reversed in diabetics compared to non-diabetics. That is, the directed alternative captures especially proteins behaving such that heparin either increases protein levels in non-diabetics but there is less of an increase or even a descent in diabetic patients, or such that heparin decreases protein levels in non-diabetics but the decrease is less, absent or even an increase in diabetic patients (cf. in theory administration of heparin increases protein levels and this effect is expected to be more pronounced in none-diabetics, yet in order to not exclude relevant opposite effects both were included in the directed hypothesis).

To test the general alternative, we used Fisher's combined *p*-value[21] as test statistic, that is $T_{general} = -2\sum_i \log(p_i)$, where i ranges over the proteins of the feature set and p_i denotes the p-value for the interaction term from the univariate regression modelling of the i-th proteins. For the second alternative, we considered the number of proteins whose estimated interaction effect is directed opposite to the estimated main effect of heparin in the univariate regression model as a test statistic $T_{directed}$. The second test statistic is intended to be more tailored to the directed alternative.

Since the proteomic features in our dataset are correlated, null distributions were simulated for both test statistics. Especially, we applied parametric bootstrapping by firstly fitting the null model for each of the 273 proteins separately, determining corresponding effect estimates, calculating residuals (averaged over the five imputations) and estimating their covariance matrix Σ. (The matrix Σ is of dimension 273×273 and its elements e_{ij}, $1 \le i,j \le 273$, are the co-variances between the residuals from two proteins i and j.) Secondly, in the bootstrapping approach, we simulated new protein responses in a multivariate manner using the previously estimated subject- and protein-specific means together with residuals simulated from a centred multivariate normal distribution with the estimated covariance matrix Σ. Our bootstrapping approach from a multivariate normal distribution preserves the correlation structure of the proteomic features, and the association between any of the proteins and the covariates diabetes, heparin as well as age, sex and BMI, but any interactions between diabetes and heparin are dissolved in the bootstrapped samples. The described bootstrapping approach extends approaches for interaction testing of univariate data presented in Buzkova 2016[22] to pathway test-statistics based on interaction terms. Null distributions were based on 5000 bootstrap-repeats and *p*-values are one-sided and were calculated as proportion of bootstrap-repeats with a test-statistic larger than or equal to the observed test-statistic. In cases where the bootstrap test-statistic did not exceed the observed test-statistic for any of the bootstrap repeats, we reported $p < 1/5000 = 2^{*}10^{-4}$.

2.4.3. Multi-proteomics-interaction-modelling

Finally, we applied penalized regression methods for high-dimensional data to derive a model in which several proteins act together. Since high-dimensional protein data is inappropriate as multivariate outcome and since administration of heparin is reflected by proteins, we considered heparin as binary outcome variable. Using model selection with diabetes, age, sex and BMI, all proteins as well as all pairwise diabetes-protein interactions as potential predictors, our research question was transferred into the question, whether diabetes-protein interactions contribute to the prediction, i.e. Does diabetes change the way in which proteins predict administration of heparin?

For model selection, we used LASSO logistic regression together with stability selection as described in Meinshausen et al.[23] and the improved sampling approach as described in Shah et al.[24] Stability selection selects predictors contributing consistently in repeated subsamples and thereby controls the expected number of potential noise variables

(variables with low selection probability) chosen into the model. For each predictor, its selection frequency describes for how many subsamples a predictor was chosen, which can be used as a measure of variable importance. The variables diabetes, age, sex and BMI were forced into all models and did not undergo selection.

To incorporate the five different imputations, we applied stability selection to each imputation separately and used the average of the five selection frequencies. We further used the following parameters: 50 random complementary pairs of subsamples (which is default in stability selection), a cut-off of 0.6 for the selection frequency, and an upper bound for the expected number of low selection probability variables of 1. These values correspond to an average of 14 selected variables for the subsample-individual models.

After model selection, we fitted an ordinary logistic regression model using the selected predictors and reported the corresponding area under the curve (AUC) as a performance measure.

All analyses were done in R version 3.6.0 together with the packages Amelia version 1.7.5, mice version 3.6.0, stabs version 0.6–3, glmnet version 2.0–18, pROC version 1.14.0, interplot version 0.2.2.[25–31]

3. Results

3.1. Results of univariate analysis

The clinical characteristics of the 497 study patients are displayed in Table 1. The results of the regression analysis of all 273 proteins can be found in Supplementary Table 4: none of the proteins showed a significant interaction between heparin and diabetes after correction for multiple testing (i.e. FDR_interact >0.05).

Table 2 contains a list of the 20 proteins with the lowest interaction *p*-values; and as seen, seven proteins displayed a non-significant *p* value for the interaction term <0.05: lipoprotein lipase, protein AMBP, complement C4—B, agrin, histidine-rich glycoprotein, kininogen-1 and isoform 2 of SPARC-related modular calcium-binding protein 1. Plots of the marginal effects of these 7 proteins are displayed in Supplementary Fig. 1 of the Supplementary Material.

3.2. Set-based analyses

For the set-based analysis, we obtained $T_{general} = 96.2$ ($P < 2^{*}10^{-4}$, i.e. none of the 5000 simulation reached a test-statistic at least this high) and $T_{directed} = 32$ ($P = 0.11$) for the set of the 44 proteins identified via the Gene Ontology database, as well as $T_{general} = 92.6$ ($P < 2^{*}10^{-4}$) and $T_{directed} = 31$ ($P = 0.17$) for the 42 proteins from the previous study of (Beck et al. 2018).[10]

3.3. Multi-proteomics-interaction-modelling

The model selection procedure yielded the following five proteins to consistently predict, whether heparin had been administered to a patient: C—C motif chemokine 5, C—C motif chemokine 28, follistatin, midkine and pleiotrophin (see Supplementary Table 5 for details). The model achieved an AUC of 0.92 in the same dataset corresponding to an excellent prediction performance of heparin. No protein-diabetes interaction was consistently contributing to the prediction of heparin, that is, diabetes did not modify the proteins' effects in the prediction model.

4. Discussion

In this study, we aimed at the identification of proteins, whose association with diabetes status and heparin treatment exhibit an interaction, i.e. proteins for which the release from the endothelial glycocalyx due to heparin treatment is changed by diabetes status. Such proteins could be highly relevant as candidates for diabetes related endothelial glycocalyx dysfunction. We do, however, not observe individual proteins with significant interaction (i.e. with FDR < 0.05) in the univariate analyses after taking multiple testing into account (see Table 2), nor identify proteins for which the interaction with diabetes predicts heparin treatment in a composite manner in the multi-proteomics-interaction analysis (see Supplementary Table 5). We do, nonetheless, find consistent interactions in the general set-based analyses (i.e. $p < 2^{*}10^{-4}$), indicating that some of the heparin-binding proteins identified via the GO database, as well as some of the proteins identified to change in level after heparin treatment in the previous study of Beck et al. 2018,[10] do respond differently to heparin administration with diabetes, supporting our hypothesis and potentially shedding light onto new mechanisms behind endothelial glycocalyx dysfunction. However, as the directed set-based analysis was not significant ($p > 0.05$), we cannot derive the specific direction of effect of these interactions in the present study.

Several of the seven proteins observed to have nominally significant interactions (i.e. P_interact<0.05, see Table 2) in the univariate analysis

Table 2
Top 20 proteins from univariate analysis.

Protein identifier	Protein name	diabetes	heparin	diabetes.heparin	P_diab_total	P_hep_total	P_interact	FDR_interact
P06858	Lipoprotein lipase	0.2	1.3	−0.58	7.18e-03	3.88e-48	7.99e-03	0.99
P02760	Protein AMBP	0.79	0.25	−0.72	2.73e-03	1.37e-02	8.64e-03	0.99
P0C0L5	Complement C4-B	0.43	−0.27	−0.69	4.66e-02	6.18e-05	1.10e-02	0.99
O00468	Agrin	−0.22	0.12	0.6	3.42e-02	5.62e-03	2.57e-02	0.99
P04196	Histidine-rich glycoprotein	0.16	0.79	−0.58	1.47e-02	4.18e-12	2.81e-02	0.99
P01042	Kininogen-1	0.078	0.94	−0.53	6.84e-03	1.71e-19	3.17e-02	0.99
Q9H4F8_2	Isoform 2 of SPARC-related modular calcium-binding protein 1	0.3	1	−0.5	1.17e-01	5.34e-23	3.91e-02	0.99
P01860	Ig gamma-3 chain C region	−0.48	−0.28	0.58	1.08e-01	3.24e-02	5.04e-02	0.99
P06316	Ig lambda chain V—I region BL2	0.1	−0.25	−0.52	3.16e-02	4.02e-04	5.50e-02	0.99
Q7Z5L7_3	Isoform 3 of Podocan	0.26	0.75	−0.49	1.45e-01	3.05e-11	5.87e-02	0.99
P23083	Ig heavy chain V—I region V35	−0.39	−0.32	0.49	1.75e-01	1.53e-02	6.45e-02	0.99
O15230	Laminin subunit alpha-5	0.35	0.84	−0.51	2.27e-01	1.05e-09	6.46e-02	0.99
P21246	Pleiotrophin	0.14	1.5	−0.37	1.33e-01	1.64e-64	8.23e-02	0.99
O95715	C-X-C motif chemokine 14	0.19	1.2	−0.42	1.74e-01	7.77e-28	8.75e-02	0.99
P01857	Ig gamma-1 chain C region	0.45	−0.33	−0.45	1.34e-01	5.56e-05	1.00e-01	0.99
Q99879	Histone H2B type 1-M	−0.17	−0.51	0.45	1.56e-01	5.06e-05	1.04e-01	0.99
P01008	Antithrombin-III	0.2	1.3	−0.36	2.66e-01	1.56e-48	1.08e-01	0.99
P06310	Ig kappa chain V-II region RPMI 6410	0.24	−0.27	−0.42	2.71e-01	9.15e-04	1.16e-01	0.99
P10646	Tissue factor pathway inhibitor	0.18	1.3	−0.35	2.83e-01	4.85e-45	1.20e-01	0.99
Q9H772	Gremlin-2	0.25	0.94	−0.42	2.69e-01	1.08e-15	1.23e-01	0.99

Note: diabetes, heparin, diabetes_heparin are the estimated effects of diabetes, heparin and the interaction of both from the protein-wise regression model with log-transformed and standardized protein levels as outcome. P_interact and FDR_interact correspond to the significance of the interaction effect. P_diab_total and P_hep_total are p-values from the Wald tests of the two joint hypotheses H_0: diabetes = diabetes_heparin = 0 (no effect of diabetes) and H_0: heparin = diabetes_heparin = 0 (no effect of diabetes).

are well-known plasma proteins, which are directly relevant for endothelium or glycocalyx function or heparin biology. Lipoprotein lipase (LPL) plays among others a key role in lipid clearance from the blood stream and is known to be recruited to the luminal surface of the vascular endothelium by heparan sulfate proteoglycans.[32] Histidine-rich glycoprotein (HRG) is a plasma glycoprotein with diverse functions including coagulation, fibrinolysis, cell chemotaxis, cell adhesion, angiogenesis and immune complex and pathogen clearance.[33,34] The activity of HRG is regulated by heparin and heparan sulfate and of specific relevance for endothelial cells HRG has protective effects on vascular barrier function.[35] Furthermore, the agrin protein is a heparan sulphate basal lamina glycoprotein, which is involved in the formation and the maintenance of neuromuscular junctions especially relevant for postsynaptic differentiation, and its isoforms 3 and 6 are muscle specific and involved in endothelial cell differentiation.[34] Lastly, two of the seven proteins have been reported to affect vascular permeability; high molecular weight kininogen, an isoform of kininogen-1, and the C4a anaphylatoxin, derived from proteolytic degradation of complement C4.[36,37] Hence, these proteins appear to relate different aspects of vascular endothelium function and heparin biology. Performing a post hoc gene enrichment analysis with the MSigDB tool (http://www.gsea-msigdb.org) revealed the gene set 'Complements and coagulation cascades' (KEGG hsa04610) to be significantly enriched (FDR corrected p-value <0.05), hence supporting a role in vascular function. Furthermore, of the seven nominal significant proteins, kininogen-1, complement C4—B and LPL have been reported to associate to diabetes,[38–40] while kininogen-1 and agrin have been linked to diabetic nephropathy[41,42] and protein AMBP and LPL have been linked to diabetic albuminuria[43] and diabetic dyslipidaemia,[44] respectively. Thus, the indications, that these heparin-influenced proteins may be of interest in relation to diabetes make sense from a biological point of view and the lack of significant findings in our univariate analysis in the present study may simply be due to lack of statistical power, as both proteome-wide (multiple) hypothesis testing, as well as testing effect modification (i.e. interactions) require a large sample sizes.

Interestingly, out of the seven proteins, which were nominal significant in the univariate analysis, LPL, HGR, agrin and kininogen-1 were also identified in the GO database search of the present study, and LPL, HGR and kininogen-1 were found to change in level after heparin treatment in the previous study by Beck et al.[10] Consequently, these proteins were among the proteins investigated in the set-based analyses. Furthermore, both sets of heparin-relevant proteins investigated in the set-based analyses also included amine oxidase, midkine, antithrombin-III, complement factor H, follistatin, pleiotrophin and secreted frizzled-related protein 1. These seven proteins also revealed the 'Complements and coagulation cascades' as significant (FDR corrected p-value <0.05), therefore, reflecting a role in cardiovascular function. Moreover, of these proteins follistatin, an inhibitor of follicle-stimulating hormone production and release, has been linked to diabetes,[45] while pleiotrophin, a heparin binding protein affecting among others processes like endothelial cell migration and cell growth and survival, has been reported for diabetic retinopathy.[46] Moreover, complement factor H, a glycoprotein essential for immune response via its role in complement activation, has been associated to insulin resistance[47] and antithrombin-III, a plasma protase inhibitor involved in regulation of the blood coagulation cascade, has been linked to diabetic retinopathy or nephropathy.[48] These proteins did together with the remaining heparin-relevant proteins (see Supplementary Tables 2 and 3) reveal significance in the general set-based analysis, i.e. the analysis reflecting interaction between diabetes and heparin for at least one protein. Yet, the directed set-based analysis, which reflected that diabetes weakens or reverses the effect of heparin on protein and hence theoretically best reflects glycocalyx endothelia dysfunction, did not show significant association (i.e. $p > 0.05$). Hence, the set-based analyses of the present study indicate that the proteins of the investigated sets are relevant for endothelia dysfunction, yet the specific direction of effect cannot be pinpointed. This finding likely indicates that the effect of the heparin and diabetes interaction relative to the individual proteins, contributing to the combined sets, might be of different magnitudes and directions and that the protein-basis of endothelial dysfunction is likely complex.

The strengths of the present study are first of all the statistical approach deciphering the mediating influence of diabetes on the heparin effect on protein levels, as well as the proteome-wide data enabling a hypothesis-free identification of novel protein candidates. One limitation is, however, that the present dataset does not hold detailed information on when the heparin treatment was given. Consequently, we might have some variation in the time from heparin treatment to blood sampling and as this time duration is known to affect the level of plasma proteins,[10] this might have influenced the results obtained. The dataset does, however, hold information on time from diagnosis to coronary angiography, which likely could reflect time from heparin treatment to blood sampling in most cases. In general, the distribution of this time span was rather similar for the majority of individuals, whom received heparin: 95% of the individuals who received heparin had a time from diagnosis to coronary angiography below 3.2 h (data not shown). Finally, one might also speculate that the dose of heparin given could affect the results obtained in the present study. Yet, the patients of the study population were given a standard dose of 5000 U heparin, hence the heparin dose mostly likely does not influence the results obtained. Furthermore, the present study population is composed of diabetes patients who were all admitted to the hospital due to acute coronary angiography. Moreover, the large part of the patients were diabetes type 2 patients (92%). Therefore, the results obtained here might not be generalizable to all diabetic patients and future studies applying the present statistical models in such cohort would be favourable. Finally, one could speculate that in addition to exploration of diabetes-related endothelial dysfunction, as investigated in the present study, analyses of additional phenotypes of relevance to vascular endothelial dysfunction in general, such as dyslipidemia, could bring forward potential protein biomarkers. Such analyses are, however, out of scope given our focus on diabetes-related endothelial dysfunction and would furthermore, increase the number of statistical tests and potentially impose a power problem. Nonetheless, future studies applying the present type of statistical approach to such phenotypes in other study populations are warranted.

In conclusion, the findings are compatible with the notion that diabetes interacts with the effect of heparin on the plasma proteome, probably due to glycocalyx alterations. Moreover, the present study points to several proteins as potential candidates for such endothelial glycocalyx dysfunction, including proteins related to vascular function. Furthermore, the study points to the use of statistical interaction models to decipher complex biological interactions, which could result in important new knowledge regarding the biology of endothelial dysfunction in diabetes. Nonetheless, additional studies are needed to validate and confirm these findings, and to further explore which individual protein might be candidates as a novel biomarker. In any case, the present study brings promise for future studies of protein biomarkers of endothelial dysfunction.

CRediT authorship contribution statement

MS and BD: conception and design of the study, performance of data and bioinformatics analyses, interpretation of results, and writing of the original draft.

UH and JH: conception and design of the study, interpretation of results, and review and editing of manuscript draft.

JEM, CH and MF: data acquisition, funding, interpretation of results, and review and editing of manuscript draft.

HCB and LMR: conception and design of the study, data acquisition, funding, interpretation of results and review and editing of manuscript draft.

All authors: read and approved the final version of the manuscript.

Acknowledgements

This work was supported by the Danish Heart Foundation, Odense University Hospital Research Fund (Grant R22-A1187-B615) and the Lundbeck Foundation.

Appendix A. Supplementary data

The Supplementary_Materials_and_Results.excl contains additional clinical information regarding the study population (Supplementary Table 1), as well as the two sets of proteins investigated in the set-based analyses (Supplementary Tables 2 and 3), the findings for the individual proteins in the regression analysis (Supplementary Table 4) and the multi-proteomics-interaction analysis (Supplementary Table 5) and the figures of the marginal effects of diabetes for proteins with P-interaction<0.05 (Supplementary Figure 1). Supplementary data to this article can be found online at https://doi.org/10.1016/j.jdiacomp.2021.107906.

References

1. Sena CM, Pereira AM, Seiça R. Endothelial dysfunction- a major mediator of diabetic vascular disease. Biochim Biophys Acta 2013;1832:2216-31. https://doi.org/10.1016/j.bbadis.2013.08.006.
2. Cao R-N, Tang L, Xia Z-Y, Xia R. Endothelial glycocalyx as a potential therapeutic target in organ injuries. Chin Med J (Engl) 2019;132:963-75. https://doi.org/10.1097/CM9.0000000000000177.
3. Kolálová H, Barbora Ambruzová B, Šindlerová LS, Klinke A, Kubala L. Modulation of endothelial glycocalyx structure under inflammatory conditions. Mediators Inflamm 2014;2014, 694312. https://doi.org/10.1155/2014/694312.
4. Lemkes BA, Nieuwdorp M, Hoekstra BJL, Holleman F. The glycocalyx and cardiovascular disease in diabetes: should we judge the endothelium by its cover? Diabetes Technol Ther 2012;14:S3-10. https://doi.org/10.1089/dia.2012.0011.
5. Spiess BD. Heparin: effects upon the glycocalyx and endothelial cells. J Extra Corpor Technol 2017;49:192-7.
6. Yilmaz O, Afsar B, Ortiz A, Kanbay M. The role of endothelial glycocalyx in health and disease. Clin Kidney J 2019;12:611-9. https://doi.org/10.1093/ckj/sfz042.
7. Deckert T, Feldt-Rasmussen B, Borch-Johnsen K, Jensen T, Kofoed-Enevoldsen A. Albuminuria reflects widespread vascular damage. The Steno hypothesis Diabetologia 1989;32:219-26. https://doi.org/10.1007/BF00285287.
8. Goligorsky MS. Vascular endothelium in diabetes. Am J Physiol Renal Physiol 2017;312:F266-75. https://doi.org/10.1152/ajprenal.00473.2016.
9. Wang S, Hu S, Zhong M. Proteomic investigation of the severe preeclampsia treatment by low molecular weight heparin. Clin Exp Obstet Gynecol 2014;41:620-6.
10. Beck HC, Jensen LO, Gils C, Ilondo AMM, Frydland M, Hassager C, et al. Proteomic discovery and validation of the confounding effect of heparin administration on the analysis of candidate cardiovascular biomarkers. Clin Chem 2018;64:1474-84. https://doi.org/10.1373/clinchem.2017.282665.
11. Myrup B, Yokoyama H, Kristiansen OP, Østergaard PB, Olivecrona T. Release of endothelium-associated proteins into blood by injection of heparin in normal subjects and in patients with Type 1 diabetes. Diabet Med 2004;21:1135-40. https://doi.org/10.1111/j.1464-5491.2004.01313.x.
12. Obling L, Frydland M, Hansen R, Moller-Helgestad OK, Lindholm MG, Holmvang L, et al. Risk factors of late cardiogenic shock and mortality in ST-segment elevation myocardial infarction patients. Eur Heart J Acute Cardiovasc Care 2018;7:7-15. https://doi.org/10.1177/2048872617706503.
13. http://proteomecentral.proteomexchange.org/cgi/GetDataset?ID=PXD008468.
14. The Gene Ontology Ressource. http://geneontology.org. Accessed on the 4 November 2019.
15. Gene Ontology and GO Annotation. https://www.ebi.ac.uk/QuickGO. Accessed on the 4 November 2019.
16. STRING (Protein-Protein Interaction Networks Functional Enrichment Analysis). https://string-db.org. Accessed on the 4 November 2019.
17. Benjamini Y, Hochberg Y. Controlling the false discovery rate a . practical and powerful approach to multiple testing. J R Stat Soc Ser B Methodol 1995;57:289-300. https://www.jstor.org/stable/2346101?seq=1.
18. Rubin DB. Multiple imputation for nonresponse in surveys. Wiley series in probability and statistics. NJ, USA: John Wiley & Sons, Inc., Hoboken; 1987.
19. Wang K, Li M, Hakonarson H. Analysing biological pathways in genome-wide association studies. Nat Rev Genet 2010;11:843-54. https://doi.org/10.1038/nrg2884.
20. Holmans P. Statistical methods for pathway analysis of genome-wide data for association with complex genetic traits. Adv Genet 2010;72:141-79. https://doi.org/10.1016/B978-0-12-380862-2.00007-2.
21. Fisher RA. Statistical Methods for Research Workers. Interaction. , 5th edEdinburgh: Oliver and Boyd Buzkova P; 2016. p. 22119-28.
22. Buzkova P. Interaction testing: residuals-based permutations and parametric bootstrap in continuous, count, and binary data. Epidemiol Methods 2016;5:119-28. https://doi.org/10.1515/em-2015-0010.
23. Meinshausen N, Bühlmann P. Stability selection. J R Stat Soc Ser B 2010;72:417-73. https://doi.org/10.1111/j.1467-9868.2010.00740.x.
24. Shah RD, Samworth RJ. Variable selection with error con, trol: another look at stability selection. J R Stat Soc Ser B Stat Methodol 2013;75:55-80, https://www.jstor.com/stable/23361014.
25. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org.
26. Honaker J, King G, Blackwell M. Amelia II: a program for missing data. J Stat Softw 2011;45:1-47. https://doi.org/10.18637/jss.v045.i07.
27. van Buuren S, Groothuis-Oudshoornmice K. Multivariate imputation by chained equations in R. J Stat Softw 2011;45:1-67. https://doi.org/10.18637/jss.v045.i03.
28. Hofner B, Boccuto L, Goeker M. Controlling false discoveries in high-dimensional situations: boosting with stability selection. BMC Bioinformatics 2015;16:144. https://doi.org/10.1186/s1259-015-0575-3.
29. Friedman J, Tibshirani HT. Regularization paths for generalized linear models via coordinate descent. J Stat Softw 2010;33:1-22. https://doi.org/10.18637/jss.v033.i01.
30. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez J-C, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics 2011. https://doi.org/10.1186/1471-2105-12-77. 12:77.
31. Solt F, Yueinterplot H. https://CRAN.R-project.org/package=interplot.
32. Lutz EP, Merkel M, Kako Y, Melford K, Radner H, Breslow JL, et al. Heparin-binding defective lipoprotein lipase is unstable and causes abnormalities in lipid delivery to tissues. J Clin Invest 2001;107:1183-92. https://doi.org/10.1172/JCI11774.
33. Jones AL, Hulett MD, Parish CR. Histidine-rich glycoprotein: a novel adaptor protein in plasma that modulates the immune, vascular and coagulation systems. Immunol Cell Biol 2005;83:106-18. https://doi.org/10.1111/j.1440-1711.2005.01320.x.
34. Gene Cards – The human gene database. https://www.genecards.org/. Accessed 18 June 2020.
35. Gao S, Wake H, Gao Y, Wang D, Mori S, Liu K, et al. Histidine-rich glycoprotein ameliorates endothelial barrier dysfunction through regulation of NF-κB and MAPK signal pathway. Br J Pharmacol 2019;176:2808-24. https://doi.org/10.1111/bph.14711.
36. Wang H, Ricklin D, Lambris JD. Complement-activation fragment C4a mediates effector functions by binding as untethered agonist to protease-activated receptors 1 and 4. Proc Natl Acad Sci U S A 2017;114:10948-53. https://doi.org/10.1073/pnas.1707364114.
37. Fernando LP, Natesan S, Joseph K, Kaplan AP. High molecular weight kininogen and factor XII binding to endothelial cells and astrocytes. Thromb Haemost 2003;90:787-95. https://doi.org/10.1160/TH03-04-0231.
38. Meng Q, Ge S, Yan W, et al. Screening for potential serum-based proteomic biomarkers for human type 2 diabetes mellitus using MALDI-TOF MS. Proteomics Clin Appl 2017;11. https://doi.org/10.1002/prca.201600079.
39. Lepedda AJ, Lobina O, Rocchiccioli S, Nieddu G, Ucciferri N, De Muro P, et al. Identification of differentially expressed plasma proteins in atherosclerotic patients with type 2 diabetes. J Diabetes Complications 2016;30:880-6. https://doi.org/10.1016/j.jdiacomp.2016.03.007.
40. Hatefi Z, Soltani G, Khosravi S, Kazemi M, Salehi AR, Salehi R. Micro R-410 binding site single nucleotide polymorphism rs13702 in lipoprotein lipase gene is effective to increase susceptibility to type 2 diabetes in Iranian population. Adv Biomed Res 2018 May 23;7:79. https://doi.org/10.4103/abr.abr_286_16.
41. Vitova L, Tuma Z, Moravec J, Kvapil M, Matejovic M, Mares J. Early urinary biomarkers of diabetic nephropathy in type 1 diabetes mellitus show involvement of kallikrein-kinin system. BMC Nephrol 2017;18:112. https://doi.org/10.1186/s12882-017-0519-4.
42. Devetzis V, Daryadel A, Roumeliotis S, Theodoridis M, Wagner CA, Hettwer S. Huynh-Do U1, Ploumis P, Arampatzis S. C-Terminal fragment of Agrin (CAF): a novel marker for progression of kidney disease in type 2 diabetics. PLoS One 2015;10, e0143524. https://doi.org/10.1371/journal.pone.0143524.
43. Shao B, Zelnick LR, Wimberger J, et al. Albuminuria, the high-density lipoprotein proteome, and coronary artery calcification in type 1 diabetes mellitus. Arterioscler Thromb Vasc Biol 2019;39:1483-91.
44. Mead JR, Irvine SA, Ramji DP. Lipoprotein lipase: structure, function, regulation and role in disease. J Mol Med (Berl) 2002;80:753-69. https://doi.org/10.1007/s00109-002-0384-9.
45. Sylow L, Vind BF, Kruse R, Møller PM, Wojtaszewski JFP, Richter EA, et al. Circulating follistatin and activin A and their regulation by insulin in obesity and type 2 diabetes. J Clin Endocrinol Metab 2020;105. https://doi.org/10.1210/clinem/dgaa090. pii: dgaa090.
46. Zhu X, Bai Y, Yu W, Pan C, Jin E, Song D, Xu Q, Yao Y, Huang L, Tao Y, Li X, Zhao M. The effects of pleiotrophin in proliferative diabetic retinopathy. PLoS One 2015, 24;10(1):e0115523. https://doi.org/10.1371/journal.pone.0115523
47. Moreno-Navarrete JM, Martínez-Barricarte R, Catalán V, Sabater M, Gómez-Ambrosi J, Ortega FJ, et al. Complement factor H is expressed in adipose tissue in association with insulin resistance. Diabetes 2010;59:200-9. https://doi.org/10.2337/db09-0700.
48. Asakawa H, Tokunaga K, Kawakami F. Elevation of fibrinogen and thrombin-antithrombin III complex levels of type 2 diabetes mellitus patients with retinopathy and nephropathy. J Diabetes Complications 2000;14:121-6. https://doi.org/10.1016/s1056-8727(00)00075-1.