Supplementary Material

Index

I. Supplementary Methods

Study participants (DapKid: A randomised, placebo-controlled, double-blinded crossover trial)

The current discovery study is based on a randomised, placebo-controlled, double-blinded crossover trial coined DapKid, which was carried out to investigate the changes dapagliflozin induces on urinary proteomics. The study has been registered on ClinicalTrials.gov with clinical trial identification number NCT02914691. This study population (n = 40) has been previously described together with the inclusion and exclusion criteria [S1]. Briefly, the inclusion criteria were the following: 1) Type 2 diabetes (T2D) patients (WHO criteria) of age 18 years and above, 2) four weeks of stable anti-glycaemic treatment or insulin before and during the study, 3) anti-hypertensive treatment (including RAAS blockade) for four weeks and throughout the study, 4) HbA1c > 58 mmol/mol (7.5%), eGFR ≥ 45 mL/min/1.73m2 5) UACR ≥ 30 mg/g in at least two out of the three morning spot urine samples before the randomisation, 6) eGFR ≥ 45 mL/min/1.73m^2 (basis CKD-EPI 2009 equation) [S2]. In the initial version of the protocol the requirement was set to eGFR ≥ 60 mL/min/1.73m^2 but later it was amended to eGFR ≥ 45 mL/min/1.73m^2 due to difficulties in recruiting participants with the initial threshold. The exclusion criteria comprised: 1) current treatment with loop diuretics, thiazolidinediones, or SGLT2i, 2) cancer treatment, 3) severe hepatic insufficiency or abnormal liver function, 4) pregnancy or breastfeeding, 5) patients at risk for dehydration, 6) congestive heart failure (New York Heart Association class IV), 7) total bilirubin > 2.0 mg/dL (34.2 μmol/L), 8) unstable or rapidly progressing renal disease, and 9) volume depletion. Additionally, patients with a recent cardiovascular (CV) event (within two months before the start of the study), including Acute Coronary Syndrome, hospitalisation for unstable angina or myocardial infarction, acute stroke or transient ischaemic attack, or recent post-coronary artery re-vascularisation, were excluded from the study. The participants included in the study were recruited at the outpatient clinic of Steno Diabetes Center Copenhagen (SDCC), Denmark between August 2015 to February 2017. After screening 40 participants were randomised in a 1:1 ratio into two sequence groups (Dapagliflozin 10 mg once daily or placebo). During the active medication period, participants received 10 mg of dapagliflozin (provided by AstraZeneca, Södertälje, Sweden) orally for 12 weeks. Dapagliflozin treatment was matched with an equal portion of placebo for another 12 weeks. Overall duration being 24 weeks (2 × 12 weeks). All participants received stable RAAS blockade treatment throughout the study. The trial was completed in July 2017.

The study design is shown in figure S1. The two equal-sized groups received the treatment in opposite orders. and no washout period was included in the study when the two groups crossed over. The clinical factors and urinary proteomics were measured at baseline (visit 1 in figure S1) and the end of both placebo and dapagliflozin treatment periods (visits 2 and 3 in figure S1), thus producing three data points for each participant.

Primary and Secondary outcomes

Primary outcome

Urinary peptide (fragment abundance) changes before and after dapagliflozin treatment between the intervention groups were assessed as a primary outcome in DapKid study.

Secondary outcomes

Secondary post hoc analyses included examining changes in ten clinical factors namely, weight (kg), body mass index or BMI (kg/m^2), HbA_{1c} (mmol/mol), urinary albumin creatinine ratio or UACR (mg/g), ambulatory systolic blood pressure (mmHg), diastolic blood pressure (mmHg), serum creatinine (µmol/L), mGFR ($mL/min./1.73m^2$) (51Cr-EDTA). LDL cholesterol (mmol/L) and ALAT (U/L) before and after treatment and comparing between groups.

Randomization

Patients were randomly assigned to a sequence of treatment (dapagliflozin followed by placebo or vice versa) in clusters of four (e.g. 2 starting with placebo followed by 2 with dapagliflozin). Allocation sequence was generated by The Capital Region of Copenhagen (RegionH) pharmacy and assigned participants to intervention while the study investigators enrolled the participants for randomization. The randomization list was generated by a validated system that automatically performs the random assignment of patients to randomization numbers. The randomization list was kept from the patients and investigators (double blinded) and the randomization code was only broken upon trial completion.

Sample size and power

As this study (Randomised cross over trial with proteomics) is the first of its kind, and of exploratory nature, no power calculation was carried out while designing. Previous experience on clinical studies

indicated that a crossover design will need about 20 to 60 patients for reasonable power. 40 subjects are reasonable for a 2×2 crossover pharmacodynamics study. With pharmacological effects on continuous endpoints (proteomics or clinical factors), online calculators (https://clincalc.com/stats/samplesize.aspx and http://hedwig.mgh.harvard.edu/sample_size/size.html) suggested a power of ≥ 80% for a mean peptide abundance fold change of 50% between groups with alpha 0.05, enrolment ratio of 1 and study size of 40 participants.

Urine Proteomics Measurement

Sample preparation and CE-MS analysis

Urine aliquots were thawed and 700 µl mixed with 700 µl of 2 M urea, 10 mM NH4OH containing 0·02 % SDS. Subsequently, samples were ultrafiltered using a Centristat 20 kDa cut-off centrifugal filter device (Satorius, Göttingen, Germany) to eliminate high molecular weight proteins. The obtained filtrate was desalted using a PD 10 gel filtration column (GE Healthcare Bio Sciences, Uppsala, Sweden) to remove urea, electrolytes, and salts as well as to enrich polypeptides. The samples were lyophilized and stored at 4°C until usage. Shortly before CE-MS analysis, the samples were re-suspended in 10 µl HPLC-grade water. Samples were injected into CE-MS with 2 psi for 99 seconds resulting in injection volumes of ~280 nl.

A P/ACE MDQ capillary electrophoresis system (Beckman Coulter, Fullerton. CA) was coupled with a MicrOTOF II MS (Bruker Daltronic, Bremen, Germany). A solution of 20% acetonitrile (Sigma-Aldrich, Taufkirchen, Germany) in HPLC-grade water (Roth, Karlsruhe, Germany) supplemented with 0·94% formic acid (Sigma-Aldrich) was used as running buffer. For CE-MS analysis, the electrospray ionization interface from Agilent Technologies (Palo Alto, CA) was set to a potential of -4·0 to -4·5 kV. Spectra were recorded over an m/z range of 350-3000 and accumulated every 3 seconds.

CE-MS data processing

After the CE-MS analysis, mass spectral ion peaks representing identical molecules at different charge states were deconvoluted into single masses using MosaFinder software [S3]. Only signals with z>1 observed in a minimum of 3 consecutive spectra with a signal-to-noise ratio of at least 4 were considered. The resulting peak list characterizes each polypeptide by its mass and migration time. Data were calibrated utilizing 3151 internal standards as reference data points for mass and migration time by applying global and local linear regression, respectively. Reference signals of 29

abundant peptides were used as internal standards for calibration (normalization) of signal intensity using linear regression. This approach is superior to normalization using urine creatinine and has been described previously [S4]. This procedure is highly reproducible and addresses both analytical and dilution variances in a single calibration step. Among, 60 independent analytic runs of a single urine sample, the coefficient of variation of peptide panel was 1% [S5]. The obtained peak list characterizes each polypeptide by its calibrated molecular mass [Da], calibrated CE migration time [min] and normalized signal intensity. All detected peptides were deposited, matched, and annotated in a Microsoft SQL database allowing further statistical analysis.

Sequencing of peptides

Peptides were sequenced using CE-MS/MS or LC-MS/MS analysis. MS/MS experiments utilized an Ultimate 3000 nano-flow system (Dionex/LC Packings, USA) or a P/ACE MDQ capillary electrophoresis system (Beckman Coulter, Fullerton, CA), both connected a Q Exactive™ Plus Hybrid Quadrupole-Orbitrap™ Mass Spectrometer (ThermoFisher Scientific, Waltham, Massachusetts, USA). Survey full-scan MS spectra (from m/z 300–2000) were acquired in the Orbitrap and ions were sequentially isolated for fragmentation. Resulting data files were searched against the UniProt human nonredundant database using Proteome Discoverer 2.4 and SEQUEST search engine. Relevant settings included: no fixed modifications, oxidation of methionine and proline as variable modifications. Minimum precursor mass was set to 790 Da. maximum precursor mass to 6000 Da with a minimum peak count of 10. FDR was set to 1%, precursor mass tolerance was 5 ppm and fragment mass tolerance 0·05 Da. The correlation between peptide charge at the working pH of 2 and CE-migration time was utilized to minimize false-positive derivation rates and validation of obtained peptide sequences. Calculated CE-migration time of the sequenced candidate (based on its peptide sequence & number of basic amino acids) was compared to the experimental migration time.

Functional gene & pathway enrichment

The STRING database contains known and predicted physical and functional interactions between proteins using data from high-throughput lab experiments. genomic context predictions. automated text-mining, co-expression, and previous knowledge databases [S6]. In the current paper, the STRING network was produced using the default settings: medium confidence score of 0.40 with FDR stringency of 0.05, and the produced network was then edited on Cytoscape. ClueGO, a functional enrichment analysis tool on Cytoscape was also employed to see whether any term was

significantly associated to the dapagliflozin affected proteins. The genes from three gene ontology (GO) knowledgebases were used: GO Biological Pathways, GO Molecular Function, and GO Cellular Component. The analysis was run using the two-tailed hypergeometric test with BH-adjusted p-value cut-off of 0.05. GO term fusion was used to avoid redundancy in the results. and κ-score threshold was set to 0.4. The default 3-8 GO tree interval was employed, i.e. at least three genes from the input had to be associated with a term for it to be considered significantly enriched. Moreover, the term was only considered significantly enriched if the genes associated to it from the input made up at least 4% of the total number of genes associated to this gene. Another Cytoscape app. CluePedia, was used to visualize the genes that were associated with the significantly enriched terms [S7].

Additional participating cohorts

1. **Independent cohort (PROTON) [S8, S9]:** This study comprises 110 type 1 diabetes (T1D) cases with micro or macroalbuminuria referred to as DKD and 50 age and gender matched healthy controls. Originally this is a cross-sectional study with a total of 161 T1D participants with varying albuminuria (50 with normoalbuminuria: <3.39 mg/mmol OR <30 mg/24 h or mg/g, 50 with microalbuminuria: 3.39–33.79 mg/mmol OR 30–299 mg/24 h or mg/g and 61 with macroalbuminuria: ≥33.90mg/mmol OR ≥300mg/24 h or mg/g) and 50 non-diabetic healthy control volunteers. T1D individuals are >18 years of age with diabetes diagnosis as per WHO criteria. T1D individuals with non-diabetic kidney disease, renal failure (eGFR<15 ml $min^{-1}[1.73m]^{-2}$), dialysis or kidney transplantation, change in RAAS blocking treatment a month prior to study inclusion, or treatment with immunosuppressive therapy were excluded. Age and gender matched control group comprised self-reported healthy volunteers recruited by newspaper advertisement from greater Copenhagen area.None of the participants were on any prescribed medication during study inclusion. Mean HbA1c levels for healthy controls were 5.4% (36 mmol/mol). More details have been reported previously [S8, S9].
2. **PROVALID [S10]:** This is a large observational, prospective study comprising type 2 diabetes (T2D) participants from five European countries. Participants with T2D (defined by ADA guidelines or getting hypoglycemic treatment) aged between 18-75 years were recruited between 2011 and 2015 and were followed up for 5 years. 88 participants from Austrian arm participated in the current study. Details have been described previously [S10].

II. Supplementary Figures

Fig. S1. Study design [S1]. Randomized, double blind, placebo-controlled, crossover study initially comprising forty participants randomized in 1:1 ratio to oral dapagliflozin 10 mg once daily or placebo added to standard treatment as described previously and conducted at the Steno Diabetes Center Copenhagen, Denmark. Baseline measures were obtained during visit 1, while the endpoints from treatment periods (visit 2 and visit 3) were used to model the changes in clinical factors and urinary proteomics.

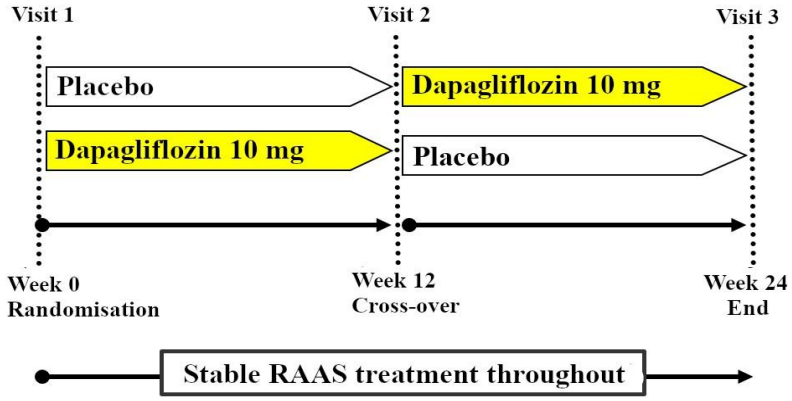


Fig. S2. Consort flow diagram. This is adapted from the first publication of this trial [S1] with participant inclusions as: presence of type 2 diabetes, HbA1c > 58 mmol/mol (7.5%), urinary albumin creatinine ratio (UACR) ≥ 30mg/g in at least 2 of 3 morning spot urine samples, eGFR ≥ 45 mL/min/1.73m^2 and all participants receiving RAAS blocking treatment on a stable dose at least prior to screening. Participants with a recent (within 2 months) cardiovascular event, congestive heart failure (HF) (NYHA IV), or unstable or acute congestive HF or on loop-diuretics were excluded. The trial began in August 2015 and ended in July 2017.

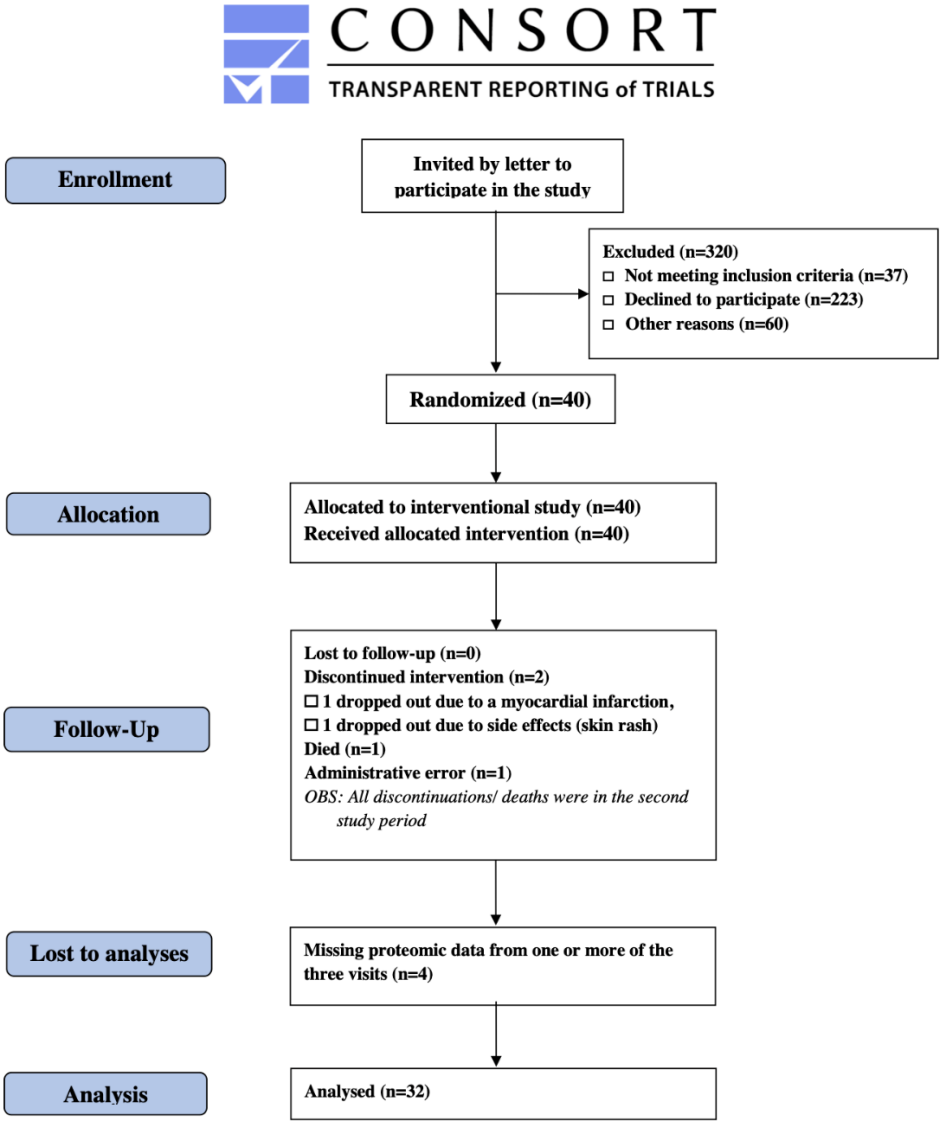


Fig. S3 Alternative output from ClueGO functional enrichment analysis. The figure shows how much proportion of the terms linked to a given term, make up, of all the terms linked to that given term. For example, from the input list of given proteins/genes, 20 were linked to collagen containing extracellular matrix which makes a little over 4% of all the terms related to this term.

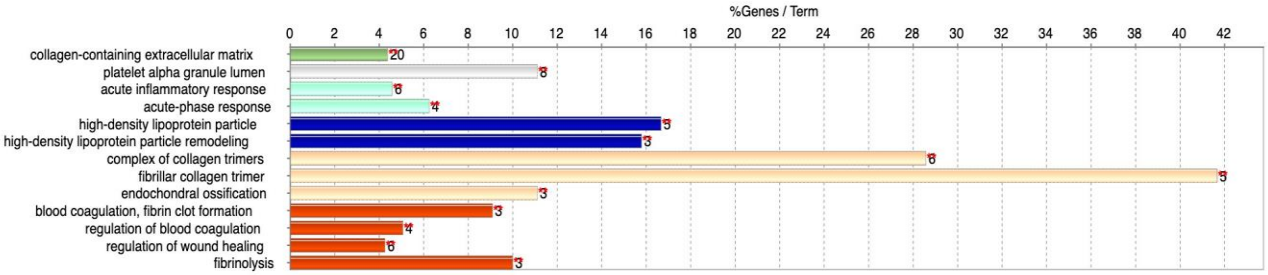
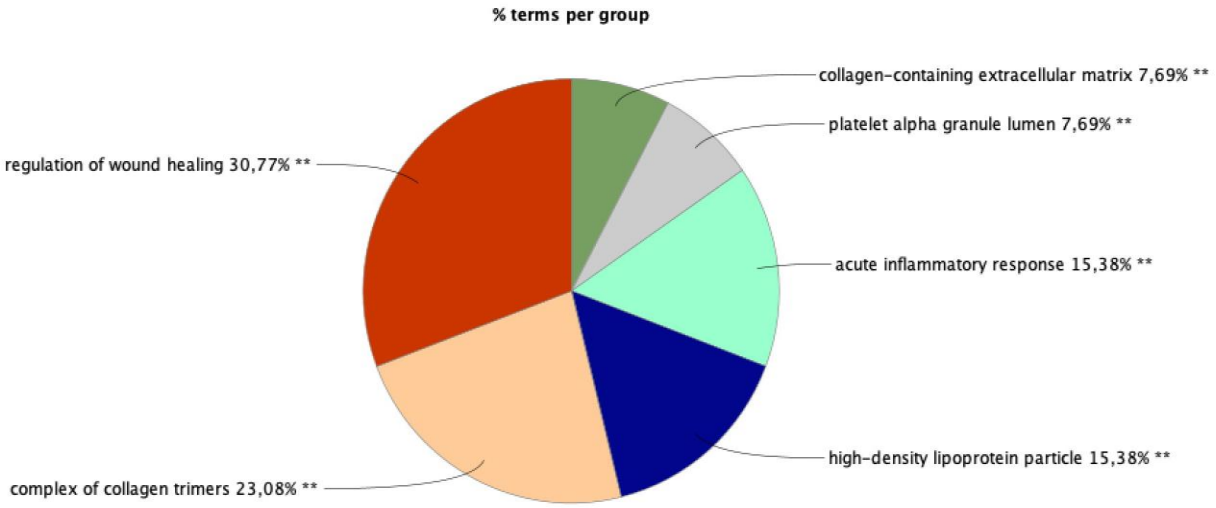


Fig. S4. ClueGO output represented as a pie chart. The figure shows an alternative representation of the output as a pie chart which shows the different categories of terms that were found to be significantly enriched. The categories are arbitrarily named after one of the terms belonging to a given class. The figure shows that wound healing related processes made the largest category of enriched terms as 30.8% (4 of the 13) significantly enriched terms belonged to this category.



III. Supplementary Tables

Supplementary Tables S1 and S2 are post hoc analyses of our previous publication where clinical changes from Dapagliflozin treatment were reported in n=35 individuals completing the trial [S1] [S1]. Here individuals with n=32 had complete proteomics data therefore these changes have been re-calculated.

Table S1. Comparison of clinical factors at baseline. The results from two-tailed Student's t-test and Mann-Whitney U test performed on the continuous clinical factors at baseline, comparing the participants from the DapKid study to the DKD cases from the independent cohort (PROTON). As UACR measurements were non-normally distributed, Mann-Whitney U test was used instead.

Clinical factor	DapKid (n = 32)	DKD (n = 110)	p-value
Age (years)	63.1 (8.3)	61.0 (9.8)	0.27
Sex (male. %)	87.5	59.1	**-**
Diabetes duration (years)	15.9 (4.7)	45.2 (13.1)	**-**
Weight (kg)	105.4 (20.2)	78.6 (17.6)	**2.5e-11**
BMI (kg/m^2)	33.7 (5.4)	26.4 (4.5)	**3.0e-12**
HbA_{1C} (mmol/mol)	72.8 (14.1)	62.3 (10.7)	**1.1e-5**
HbA_{1C} (%)	8.8 (1.2)	7.8 (1.0)	**6.5e-14**
UACR (mg/g)[a]	153.8 (94.2-328.9)	42.9 (11.0-195.8)	**8.1e-5**
Office systolic BP (mmHg)	141.1 (15.3)	135.5 (19.6)	0.13
Office diastolic BP (mmHg)	82.9 (10.2)	72.7 (9.3)	**3.3e-7**
Ambulatory systolic BP (mmHg)	146.5 (11.7)	139.2 (13.0)	**0.008**
Ambulatory diastolic BP (mmHg)	82.9 (7.9)	76.3 (6.3)	**1.0e-5**
Serum creatinine (µmol/L)	80.0 (22.3)	105.1 (49.3)	**0.005**
eGFR (mL/min/1.73m^2)[b]	85.5 (19.1)	68.5 (25.6)	**0.0006**
LDL cholesterol (mmol/L)	1.6 (0.7)	2.1 (0.7)	**0.0008**
ALAT (U/L)	40.3 (15.6)	35.6 (9.9)	**0.04**
Macroalbuminuria (%)[c]	34.4	54.5	**-**

DKD (n=110) includes T1D individuals with microalbuminuria (n=50) and macroalbuminuria (n=60). The p-value <0.05 denotes significant changes between DKD and Healthy controls (n=50) from the Proton study cohort. BMI: Body mass index; UACR: Urine Albumin Creatinine ratio; eGFR: Estimated glomerular filtration rate; HbA1c: Glycated haemoglobin; LDL: Low density lipoprotein; ALAT: Alanine Aminotransferase.

[a] Median (IQR: 25th to 75th percentile). Geometric mean of three morning visits.
[b] Estimated using 2009 CKD-EPI equation.
[c] UACR ≥ 300 mg/g

Table S2. Changes in clinical factors after dapagliflozin treatment. The investigated clinical factors described using mean and SD (median and IQR for UACR) after placebo and dapagliflozin treatment periods. The p-values are derived from the endpoint vs endpoint comparisons (n=32/group) from the two treatment periods when the data were modelled using linear mixed-effects models with patient-specific random intercepts. The p-values in bold show the clinical factors significantly improved by dapagliflozin treatment.

Parameters	Placebo	Dapagliflozin	End vs. end	CI	p-values
Weight (kg)	104.8 (19.9)	103.6 (19.8)	-2.0	-2.7 to -1.2	**< 0.00001**
Body mass index (kg/m^2)	33.6 (5.6)	33.3 (5.5)	-0.64	-0.89 to -0.39	**< 0.0001**
HbA_{1c} (mmol/mol)	73.3 (13.9)	65.6 (13.2)	-7.8	-11.2 to -4.5	**< 0.0001**
UACR (mg/g)	167.8 (75.33 – 310.0)	113.3 (67.25 – 158.33)	-33.1%	-43.2% to -21.2%	**< 0.001**
Ambulatory systolic BP (mmHg)	146.0 (10.6)	141.9 (12.2)	-4.2	-8.2 to -0.28	**0.03682**
Ambulatory diastolic BP (mmHg)	81.5 (6.0)	78.9 (6.7)	-2.7	-4.61 to -0.79	**0.007519**
Serum creatinine (µmol/L)	80.6 (22.7)	82.1 (21.1)	1.5	-2.32 to 5.37	0.4241
mGFR (mL/min/1.73m^2) (51Cr-EDTA)	88.2 (25.3)	77.4 (23.5)	-11.0	-16.3 to -5.8	**< 0.001**
LDL cholesterol (mmol/L)	1.66 (0.75)	1.67 (0.65)	0.06	-0.085 to 0.21	0.3923
ALAT (U/L)	40.8 (17.5)	41.3 (22.8)	0.45	-4.62 to 5.52	0.8569

Table S3. Detailed model statistics for changes in clinical factors after Dapaglifozin treatment (depicted in Table S2) including the sequence term. The sequence term has been denoted by DP vs PD or PD vs. DP. ”DP” signifies the sequence group receiving dapagliflozin first, while the reverse is true for sequence ”PD”. The differences of least squares means from the linear mixed-effects models based on the clinical factors from the endpoints from placebo and dapagliflozin treatment periods. The confidence intervals obtained using Kenward-Roger’s approximation for the degrees of freedom.

Clinical characteristic	Estimate	2.5% CI	97.5% CI	SE	Df	t	p-value
Weight (Kg)							
Treatment (Dapa. vs Placebo)	-1.9790	-2.7276	-1.2305	0.3659	29.0053	-5.4075	0.00001
Sequence (DP vs PD)	-5.4801	-20.006	9.046	7.1127	30.0007	-0.7704	0.447
Day (Visit 2 vs visit 3)	0.532	-0.2164	1.2805	0.3659	29.0053	1.4536	0.156
Body Mass Index (kg/m^2)							
Treatment (Dapa. vs Placebo)	-0.64418	-0.8909	-0.3974	0.12045	28.00759	-5.348	0.00001
Sequence (DP vs PD)	-1.0882	-5.2475	3.071	2.0336	29.0014	-0.5351	0.596
Day (Visit 2 vs visit 3)	0.193	-0.0536	0.4397	0.1204	28.0075	1.6026	0.12
Glycated Hemoglobin (mmol/mol)							
Treatment (Dapa. vs Placebo)	-7.8176	-11.158	-4.4769	1.6357	30.0	-4.7792	0.00004
Sequence (DP vs PD)	-3.9784	-13.235	5.27851	4.5326	30.000	-0.8777	0.387
Day (Visit 2 vs visit 3)	1.5823	-1.7583	4.923	1.6357	30.000	0.9673	0.341
UACR (mg/g)							
Treatment (Dapa. vs Placebo)	-0.3308	-0.4318	-0.2118	0.0801	30.00	-5.0121	0.00002
Sequence (DP vs PD)	0.2654	-0.4366	1.8427	0.3962	30.00	0.5941	0.556
Day (Visit 2 vs visit 3)	0.0376	-0.1190	0.2221	0.0801	30.00	0.4608	0.648
Ambulatory Systolic BP (mmHg)							
Treatment (Dapa. vs placebo)	-4.23672	-8.1917	-0.2817	1.916	23.954	-2.211	0.036
Sequence (DP vs PD)	7.30815	-0.333	14.949	3.7166	25.898	1.966	0.060
Day (Visit 2 vs visit 3)	2.41282	-1.542	6.3678	1.916	23.9547	1.259	0.22
Ambulatory Diastolic BP (mmHg)							
Treatment (Dapa. vs Placebo)	-2.6989	-4.6061	-0.7917	0.9233	23.622	-2.9231	0.0075
Sequence (DP vs PD)	1.2703	-3.4872	6.0279	2.3144	25.97245	0.5488	0.587
Day (Visit 2 vs visit 3)	0.92755	-0.9796	2.8347	0.9233	23.622	1.0046	0.325
Serum Creatinine (µmol/L)							
Treatment (Dapa. vs Placebo)	1.5274	-2.322	5.3771	1.885	30.00	0.8103	0.424
Sequence (DP vs PD)	-0.449	-16.09	15.1928	7.659	30.00	-0.0586	0.953
Day (Visit 2 vs visit 3)	-0.9392	-4.7889	2.9105	1.885	30.00	-0.498	0.621
mGFR (mL/min/1.73m^2)							
Treatment (Dapa. vs Placebo)	-11.035	-16.293	-5.777	2.574	30.00	-4.286	0.00017
Sequence (DP vs PD)	1.9686	-15.126	19.063	8.370	30.00	0.235	0.815
Day (Visit 2 vs visit 3)	3.5647	-1.693	8.8224	2.574	30.00	1.384	0.176
LDL cholesterol (mmol/L)							
Treatment (Dapa. vs Placebo)	0.0623	-0.0848	0.2095	0.0716	26.4194	0.8697	0.392
Sequence (DP vs PD)	-0.2677	-0.7792	0.2437	0.2501	28.9964	-1.0706	0.293
Day (Visit 2 vs visit 3)	0.1192	-0.0279	0.2665	0.0716	26.4194	1.6636	0.108
Alanine Aminotransferase (U/L)							

Treatment (Dapa. vs Placebo)	0.451	-4.617	5.5194	2.4789	29.227	0.1819	0.856
Sequence (DP vs PD)	3.1253	-10.847	17.098	6.8419	30.013	0.4567	0.651
Day (Visit 2 vs visit 3)	-2.8629	-7.931	2.2054	2.4789	29.227	-1.1548	0.257

CI: confidence interval, SE: standard error, t: t test statistic, Df: degrees of freedom

Table S4. Events during intervention.

Event type	**Placebo**	**Active**
Atrial fibrillation	1	1
Cardiac arrest – death	1	0
Elevated Troponin I – CABG + new aorta valve	0	1
Pneumothorax	0	1
Collapsed hip prosthesis	0	1
Psychiatric admission	0	1
Genital infection	3	6
Increased diuresis	1	5
Urinary tract infection	1	1
Genital pain	0	2
Fever	1	1
Other infections	4	2
Hypoglycemia	1	2
Fatigue	1	0
Rash	0	1
Cardiac decompensation	1	0
Dizziness	0	1
Foot Pain	1	0
Back pain	0	1
Cough	0	1
Acute kidney injury	0	0
Hypovolaemia	0	0

Table S5. Detailed changes in individual peptide fragment abundance after SGLT2i intervention on the Randomized study DapKid.

Protein	Peptide sequence ID	Placebo	Dapagliflozin	Fold Change	n	T	CI (2.5%)	CI (97.5%)	p
A1BG	11163	257.8341	19.5016	0.0756	16	136	60.6901	628.34	0.0189
A1BG	12127	3617.333	758.0441	0.2096	24	266	813.6701	4075.16	0.0271
A1BG	12950	1196.521	306.2672	0.256	21	209	370.715	1936.41	0.0307
A1BG	13296	1052.219	33.2138	0.0316	19	175	100.4101	1770.26	0.0308
AHSG	8180	355.8841	1001.749	2.8148	22	221	80.32	346.76	0.0419
ALB	12518	54307.87	803.2184	0.0148	22	251	7720.73	111157.8	0.0119
ALB	12970	91459.78	959.9853	0.0105	21	228	10228.08	151302.1	0.0124
ALB	13562	108201.1	236.7078	0.0022	24	288	1994.23	86045.16	0.0124
ALB	14191	326527	1566.885	0.0048	25	304	9579.82	541097.5	0.0138
ALB	14790	14059.24	507.8403	0.0361	19	188	4340.245	27319.07	0.0154
ALB	13726	10952.88	144.3422	0.0132	22	239	919.7401	20815.4	0.0157
ALB	8491	3482.532	39.0772	0.0112	22	236	65.595	3117.505	0.0189
ALB	8062	2556.829	19.9934	0.0078	19	177	444.9251	7281.295	0.0271
ALB	7463	9486.713	19.0097	0.002	18	157	220.67	22797.56	0.0377
ALB	5188	1971.827	52.1875	0.0265	14	101	390.88	7449.01	0.0449
APOC3	19780	178.7722	476.1769	2.6636	27	36	-416.945	-145.17	0.0157
CALD1	7837	2558.923	58.1203	0.0227	15	116	1549.595	8618.915	0.035
COL1A1	17890	590.2359	1123.486	1.9035	29	56	-888.38	-236.865	0.0189
COL1A2	19745	558.6972	1274.198	2.2807	28	52	-1140.6	-371.345	0.0201
COL3A1	13816	9461.028	16079.86	1.6996	32	52	-9213	-3887.5	0.0113
COL3A1	15237	418.1675	1003.36	2.3994	28	25	-880.62	-389.82	0.0119
COL3A1	15129	90.2166	210.5928	2.3343	25	34	-212.75	-77.355	0.0199
COL3A1	8530	490.2906	824.6206	1.6819	31	80	-497.02	-135.785	0.0271
COL3A1	11668	10751.2	14269.94	1.3273	32	98	-5023.42	-1106.53	0.0308
CRNN	883	566.4119	92.82	0.1639	18	162	149.9951	892.39	0.0271
FTSJ3	8534	207.4344	549.1441	2.6473	28	42	-525.655	-196.46	0.0157
HSP90AB1	3285	231.8928	15.1872	0.0655	22	234	57.28	441.005	0.0189
IGLL5	15015	13096.65	106.4725	0.0081	16	129	1632.035	30665.98	0.0367
KRT1	9412	185.9437	384.6772	2.0688	27	51	-372.765	-108	0.0271
PI16	8564	366.0416	66.7228	0.1823	24	276	114.98	518.465	0.0175
PIGR	18732	665.8581	1579.226	2.3717	29	42	-1375.75	-353.61	0.0138
PTGDS	8589	2263.559	148.6259	0.0657	26	295	209.53	2789.28	0.0449
SECTM1	12387	436.1622	55.0528	0.1262	20	189	122.74	880.5851	0.0369
SERPINA1	9486	9072.127	1854.2	0.2044	23	244	1255.985	8817.875	0.0308
SERPINA1	11408	4091.671	396.9144	0.097	18	157	486.16	8184.36	0.0377
TRIM33	10424	451.6884	998.2491	2.21	32	84	-566.11	-172.6	0.0189

n: individuals that had paired data from the end-to-end visits for that specific peptide; T: Test statistic for the Wilcoxon signed-rank test.

Table S6 Sequences for the urinary peptide fragments affected by dapagliflozin.

Peptide sequence ID	Protein	Gene Symbol	Peptide Sequence
11163	α-1B-glycoprotein	*A1BG*	LREGETKAVKTVRTPGAAANL
12127	α-1B-glycoprotein	*A1BG*	LREGETKAVKTVRTPGAAANLE
12950	α-1B-glycoprotein	*A1BG*	LREGETKAVKTVRTPGAAANLEL
13296	α-1B-glycoprotein	*A1BG*	VREDRGGRRVHRFQSPAGTEAL
8180	α-2-HS-glycoprotein	*AHSG*	GVVSLGSPSGEVSHPRKT
12518	Albumin	*ALB*	DAHKSEVAHRFKDLGEENFK
12970	Albumin	*ALB*	DAHKSEVAHRFKDLGEENFKA
13562	Albumin	*ALB*	DAHKSEVAHRFKDLGEENFKAL
14191	Albumin	*ALB*	DAHKSEVAHRFKDLGEENFKALV
14790	Albumin	*ALB*	DAHKSEVAHRFKDLGEENFKALVL
13726	Albumin	*ALB*	HKSEVAHRFKDLGEENFKALVL
8491	Albumin	*ALB*	LVRYTKKVPQVSTPTL
8062	Albumin	*ALB*	RFKDLGEENFKALVL
7463	Albumin	*ALB*	VRYTKKVPQVSTPTL
5188	Albumin	*ALB*	YTKKVPQVSTPTL
19780	Apolipoprotein C-III	*APOC3*	SEAEDASLLSFMQGYMKHATKTAKDALSSVQESQVAQQ
7837	Caldesmon	*CALD1*	STHQAAIVSKIDSRLE
17890	Collagen α-1(I) chain	*COL1A1*	GPpGADGQPGAKGEpGDAGAKGDAGPPGpAGPAGPPGpIG
19745	Collagen α-2(I) chain	*COL1A2*	SKGESGNKGEpGSAGPQGPpGpSGEEGKRGPNGEAGSAGPPGPpG
13816	Collagen α-1(III) chain	*COL3A1*	ERGEAGIpGVpGAKGEDGKDGSpGEpG
15237	Collagen α-1(III) chain	*COL3A1*	ERGEAGIpGVpGAKGEDGKDGSpGEpGANG
15129	Collagen α-1(III) chain	*COL3A1*	ERGEAGIpGVpGAKGEDGKDGSPGEpGANG
8530	Collagen α-1(III) chain	*COL3A1*	GARGNDGARGSDGQPGPpGP
11668	Collagen α-1(III) chain	*COL3A1*	GGpGSDGKPGppGSQGESGRPGPpG
883	Cornulin	*CRNN*	VIVKPHDPA
8534	pre-rRNA processing protein FTSJ3	*FTSJ3*	VEDDGDDTSLDSDLDPE
3285	Heat shock protein HSP 90-β	*HSP90AB1*	PEDEEEKKKM
15015	Immunoglobulin λ-like polypeptide 5	*IGLL5*	WKADGSPVKAGVETTKPSKQSNNKYA
9412	Keratin; type II cytoskeletal 1	*KRT1*	GSGGSSYGSGGGSYGSGGGGGGGRG
8564	Peptidase inhibitor 16	*PI16*	LTDEEKRLMVELHNL
18732	Polymeric immunoglobulin receptor	*PIGR*	LFAEEKAVADTRDQADGSRASVDSGSSEEQGGSSRA
8589	Prostaglandin-H2 D-isomerase	*PTGDS*	YSQGSKGPGEDFRMATL
12387	Secreted and transmembrane protein 1	*SECTM1*	HLVGHQRNNRQVTLEVSGAEP
9486	α-1-antitrypsin	*SERPINA1*	EAIPMSIPPEVKFNKPF
11408	α-1-antitrypsin	*SERPINA1*	LRTLNQPDSQLQLTTGNGLF
10424	E3 ubiquitin-protein ligase	*TRIM33*	EEDDGEVTEDSDEDFIQP

Peptide sequence ID pertains to the unique peptide ID for the current study.

Table S7. Sensitivity analysis: Detailed changes in individual peptide fragment abundance after SGLT2i intervention, independent of albuminuria lowering. Participants (n=18) with albuminuria lowering of >30% post intervention were excluded in this analysis.

Protein	Sequence ID	Placebo	Dapagliflozin	Fold Change	p-value
A1BG	12127	907.6614	252.6779	0.278	0.0185
A1BG	11163	81.7136	13.8064	0.169	0.0225
A1BG	12950	2393.3964	787.7164	0.329	0.0254
ALB	14790	110136.701	53.6464	0.0005	0.0059
ALB	13562	6130.739	37.2071	0.006	0.008
ALB	12970	5856.932	37.2364	0.006	0.0129
ALB	12518	743.9979	35.0529	0.047	0.044
ALB	8491	26105.545	303.6386	0.012	0.03
APOC3	19780	132.7893	482.7021	3.635	0.0039
COL1A1	17890	688.0929	1584.4664	2.302	0.0054
COL1A2	19745	463.085	1376.5421	2.972	0.0054
COL3A1	15237	8378.7621	14674.3357	1.751	0.0012
COL3A1	13816	52.4864	258.3736	4.922	0.0039
COL3A1	15129	335.6557	916.7157	2.731	0.0042
COL3A1	11668	9872.82	13628.925	1.380	0.0494
CRNN	883	389.6957	111.4329	0.285	0.0244
FTSJ3	8534	221.055	490.5829	2.219	0.0254
HSP90AB1	3285	168.9929	6.5257	0.038	0.0423
KRT1	9412	201.8329	357.2864	1.770	0.0455
PI16	8564	275.2064	66.7521	0.242	0.0167
PIGR	18732	750.0436	1866.8221	2.489	0.0211
SERPINA1	9486	6354.2543	2365.2721	0.372	0.0135
TRIM33	10424	590.3243	903.2029	1.53	0.0295

Table S8. Mann-Whitney U test on the urinary proteomics from the independent cohort: validation 1.

Peptide sequence ID	Gene Symbol	Protein	DKD (mean)	Healthy (mean)	p-value	Test statistic	BH-adjusted p-value
12127	*A1BG*	α-1B-glycoprotein	1781.37	4.207	0.000001	1755	**0.000016**
12950	*A1BG*	α-1B-glycoprotein	2137.423	2.469	0.000037	1902	**0.000165**
13296	*A1BG*	α-1B-glycoprotein	379.197	0	0.000264	2175	**0.000863**
11163	*A1BG*	α-1B-glycoprotein	46.823	0	0.005847	2425	**0.011694**
8180	*AHSG*	α-2-HS-glycoprotein	213.442	0	0.007883	2450	**0.014937**
12970	*ALB*	Albumin	2015.244	21.418	0.000034	1911.5	**0.000165**
13562	*ALB*	Albumin	7876.335	24.974	0.000032	1824.5	**0.000165**
12518	*ALB*	Albumin	925.61	162.918	0.000046	1907	**0.000183**
14191	*ALB*	Albumin	6433.007	12.321	0.000441	2096	**0.001322**
8491	*ALB*	Albumin	2663.851	25.193	0.023318	2453	**0.038156**
13726	*ALB*	Albumin	2867.477	4.667	0.026280	2500	**0.039290**
5188	*ALB*	Albumin	2015.093	7.35	0.039438	2501	0.052584
7463	*ALB*	Albumin	3134.85	28.44	0.046609	2452	0.059925
8062	*ALB*	Albumin	35.267	0	0.064189	2625	0.079682
14790	*ALB*	Albumin	92739.29	423.45	0.090588	2419	0.105199
19780	*APOC3*	Apolipoprotein C-III	354.975	506.002	0.022515	2207	**0.038156**
7837	*CALD1*	Caldesmon	1042.804	0	0.019239	2525	**0.034630**
17890	*COL1A1*	Collagen α-1(I) chain	960.66	1241.408	0.028376	2232	**0.039290**
19745	*COL1A2*	Collagen α-2(I) chain	1804.4	2939.349	0.000058	1703	**0.000210**
11668	*COL3A1*	Collagen α-1(III) chain	13681.574	15627.938	0.028148	2231	**0.039290**
8530	*COL3A1*	Collagen α-1(III) chain	1033.442	634.648	0.069389	2347.5	0.083266
15237	*COL3A1*	Collagen α-1(III) chain	15478.674	16458.099	0.151761	2470	0.160688
15129	*COL3A1*	Collagen α-1(III) chain	608.266	679.408	0.198791	2520	0.204471
13816	*COL3A1*	Collagen α-1(III) chain	347.915	286.865	0.405936	2685	0.405936
883	*CRNN*	Cornulin	219.596	0	0.001741	2325	**0.004477**
8534	*FTSJ3*	pre-rRNA processing protein FTSJ3	837.446	2049.9	<0.000001	1060	**<0.000001**
3285	*HSP90AB1*	Heat shock protein HSP 90-β	543.9	3.052	0.002672	2159	**0.006413**
15015	*IGLL5*	Immunoglobulin λ-like polypeptide 5	22574.673	122.896	0.127064	2560.5	0.142947
9412	*KRT1*	Keratin; type II cytoskeletal 1	564.75	1126.169	0.000003	1521	**0.000027**
8564	*PI16*	Peptidase inhibitor 16	724.295	0	0.000025	2000	**0.000165**
18732	*PIGR*	Polymeric immunoglobulin receptor	1451.439	1045.853	0.133075	2447.5	0.145173
8589	*PTGDS*	Prostaglandin-H2 D-isomerase	586.969	700.124	0.004217	2035	**0.009171**
12387	*SECTM1*	Secreted and transmembrane protein 1	234.333	17.235	0.001083	2032.5	**0.003000**
11408	*SERPINA1*	α-1-antitrypsin	1826.854	0	0.004331	2400	**0.009171**
9486	*SERPINA1*	α-1-antitrypsin	9843.728	77.306	0.026504	2336	**0.039290**
10424	*TRIM33*	E3 ubiquitin-protein ligase TRIM33	1329.369	2701.35	<0.000001	1262	**<0.000001**

Table S9. Mann-Whitney U test on the urinary proteomics from the T2D PROVALID cohort: validation 2.

Human UniProt ID	Gene Symbol	Protein	Baseline: Pre SGLT2i therapy (mean)	Post SGLT2i therapy (mean)	p-value	Start amino acid (peptide sequence)	End amino acid (peptide sequence)
P02768	*ALB*	Albumin	2933.37	68.63	0.25	25	50
	ALB	Albumin	1261.55	185.08	**3.7×10^{-5}**	433	447
	ALB	Albumin	175.24	17.42	0.13	27	48
	ALB	Albumin	0.71	0	0.21	432	444
	ALB	Albumin	473.44	53.12	**1.6×10^{-3}**	435	447
	ALB	Albumin	2.34	4.92	0.99	27	45
	ALB	Albumin	16.6	12	0.085	25	43
	ALB	Albumin	17.04	10.66	0.58	34	48
	ALB	Albumin	12555.04	732.81	**7.9×10^{-3}**	25	48
	ALB	Albumin	1012.46	284.42	0.25	25	46
	ALB	Albumin	1.73	0.95	0.75	432	449
	ALB	Albumin	628.03	153.51	**1.8×10^{-5}**	432	447
	ALB	Albumin	31.91	2.87	**9.4×10^{-3}**	436	447
	ALB	Albumin	1.4	0	0.078	31	48
	ALB	Albumin	0.14	1.2	0.55	434	447
	ALB	Albumin	1.27	1.14	0.66	30	48
	ALB	Albumin	8.84	7.88	0.60	33	48
P02452	*COL1A1*	Collagen α-1(I) chain	82.65	61.69	0.99	630	666
	COL1A1	Collagen α-1(I) chain	60.56	54.75	0.45	231	239
	COL1A1	Collagen α-1(I) chain	398.81	645.03	**2.7×10^{-4}**	221	239
	COL1A1	Collagen α-1(I) chain	184.12	215.61	0.12	231	242
	COL1A1	Collagen α-1(I) chain	441.3	566.07	**0.016**	222	241
	COL1A1	Collagen α-1(I) chain	838.4	885.25	0.20	378	399
	COL1A1	Collagen α-1(I) chain	2510.71	2832.22	0.082	1021	1041
	COL1A1	Collagen α-1(I) chain	1808.45	1971.92	0.070	231	249
	COL1A1	Collagen α-1(I) chain	5934.51	6078.48	0.12	229	249
	COL1A1	Collagen α-1(I) chain	16446.96	19383.48	**3.2×10^{-3}**	229	249
	COL1A1	Collagen α-1(I) chain	1857.17	1719.68	0.71	820	835
	COL1A1	Collagen α-1(I) chain	763.38	76.2	**0.017**	646	677
P08123	*COL1A2*	Collagen α-2(I) chain	22.48	46.46	0.59	612	625
	COL1A2	Collagen α-2(I) chain	57.1	69.13	0.50	377	389
	COL1A2	Collagen α-2(I) chain	291.18	362.83	**0.016**	920	931
	COL1A2	Collagen α-2(I) chain	28.71	38.88	0.57	46	72
P02461	*COL3A1*	Collagen α-1(III) chain	72.34	110.55	0.55	903	945
	COL3A1	Collagen α-1(III) chain	107.55	113.74	0.29	587	600
	COL3A1	Collagen α-1(III) chain	203.51	208.37	0.48	814	840
	COL3A1	Collagen α-1(III) chain	775.15	782.59	0.72	318	337
	COL3A1	Collagen α-1(III) chain	582.14	664.85	0.19	448	477
	COL3A1	Collagen α-1(III) chain	18509.89	20429.83	**0.047**	448	477
	COL3A1	Collagen α-1(III) chain	133	153.18	0.25	734	766
P01833	*PIGR*	Polymeric immunoglobulin receptor	46.21	29.4	0.34	585	604
	PIGR	Polymeric immunoglobulin receptor	123.73	38.86	0.47	604	644
	PIGR	Polymeric immunoglobulin receptor	72.81	65.69	0.39	613	648
	PIGR	Polymeric immunoglobulin receptor	951.55	920.03	0.42	599	639
P41222	*PTGDS*	Prostaglandin-H2 D-isomerase	1362.67	251	**0.038**	132	148
P01009	*SERPINA1*	α-1-antitrypsin	60.47	46.93	**6.8×10^{-3}**	167	204

Table S10. Annotation table for proteins used in functional enrichment analyses. The column ”Change”, shows whether urinary peptide fragments derived from a protein increased (+) or decreased (-) on average, and the column ”n” shows how many individual peptide fragment from a given protein were affected by dapagliflozin. If the urinary abundance of all peptide fragments derived from the same protein changed in the same direction, the number of affected peptide fragments was marked with ”*”. As the urinary level of one fragment from COL1A2 increased and one decreased, the change of this protein was marked by ±.

Gene Symbol	Change	n	Gene Name
A1BG	-	4*	α-1B-glycoprotein
AHSG	+	1	α-2-HS-glycoprotein
ALB	-	10*	Serum albumin
ANXA1	+	1	Annexin A1
APOA4	-	1	Apolipoprotein A-IV
APOC3	+	1	Apolipoprotein C-III
ARFGEF1	+	1	Brefeldin A-inhibited guanine nucleotide-exchange protein 1
B2M	-	3*	β-2-microglobulin
C3	-	1	Complement C3
CACNA1B	-	1	Calcium channel α12.2 subunit
CALD1	-	1	Caldesmon
CD99	+	1	CD99 antigen
CDH1	-	1	Cadherin-1
CLU	-	1	Clusterin
COL1A1	+	28	Collagen α-1(I) chain
COL1A2	±	2	Collagen α-2(I) chain
COL2A1	+	1	Collagen α-1(II) chain
COL3A1	+	15	Collagen α-1(III) chain
COL6A1	+	1	Collagen α-1(VI) chain
COL7A1	+	1	Collagen α-1(VII) chain
COL11A2	+	1	Collagen α-2(XI) chain
COL13A1	-	1	Collagen α-1(XIII) chain
COL23A1	+	1	Collagen α-1(XXIII) chain
CRH	+	1	Corticoliberin
CRNN	-	1	Cornulin
FGA	-	2*	Fibrinogen α chain
FGB	-	1	Fibrinogen β chain
FTSJ3	+	1	pre-rRNA processing protein FTSJ3
FXYD2	+	1	Sodium/potassium-transporting ATPase subunit γ
HSP90AB1	-	1	Heat shock protein HSP 90-β
IGF2	+	1	Insulin-like growth factor II
IGLC2	-	1	Ig λ-2 chain C regions
IGLL5	-	1	Immunoglobulin λ-like polypeptide 5
IL1RN	-	1	Interleukin-1 receptor antagonist protein
KRT1	+	1	Keratin; type II cytoskeletal 1
LEMD3	-	1	Inner nuclear membrane protein Man1
LRRC25	-	1	Leucine-rich repeat-containing protein 25
PGLYRP1	+	1	Peptidoglycan recognition protein 1
PI16	-	1	Peptidase inhibitor 16
PIGR	+	4*	Polymeric immunoglobulin receptor
PRPF3	-	1	U4/U6 small nuclear ribonucleoprotein Prp3
PTGDS	-	1	Prostaglandin-H2 D-isomerase
SAA1	-	1	Serum amyloid A-1 protein
SAA2	-	1	Serum amyloid A-2 protein
SECTM1	-	1	Secreted and transmembrane protein 1
SERPINA1	-	7*	α-1-antitrypsin
SERPINC1	-	2*	Antithrombin-III
SPATA33	+	1	Spermatogenesis-associated protein 33
SPRR3	-	1	Small proline-rich protein 3
TRIM33	+	1	E3 ubiquitin-protein ligase TRIM33
UMOD	-	1	Uromodulin
VGF	+	1	Neurosecretory protein VGF

IV. Supplementary log file from ClueGO

The log-file produced by ClueGO (v. 2.5.7.) plug-in on Cytoscape (v. 3.8.2.) is provided as text below:

ClueGO - Functional Enrichment Log File

Selection Criteria:

Statistical Test Used = Enrichment/Depletion (Two-sided hypergeometric test)

Correction Method Used = Benjamini-Hochberg

Min GO Level = 3

Max GO Level = 8

Cluster #1

Sample File Name = File selection:

Number of Genes = 3

Min Percentage = 4.0

GO Fusion = true

GO Group = true

Kappa Score Threshold = 0.4

Over View Term = SmallestPValue

Group By Kappa Statistics = true

Initial Group Size = 1

Sharing Group Percentage = 50.0

ClueGO Log:

All results were created with ClueGO v2.5.7

Organism analyzed: Homo Sapiens [9606]

Identifier types used: [SymbolID]

Evidence codes used: [All]

#Genes in GO_MolecularFunction-EBI-UniProt-GOA-ACAP-ARAP_05.05.2021_00h00 : 18336

#Genes in GO_CellularComponent-EBI-UniProt-GOA-ACAP-ARAP_05.05.2021_00h00 : 18983

#Genes in GO_BiologicalProcess-EBI-UniProt-GOA-ACAP-ARAP_05.05.2021_00h00 : 18058

#All unique genes in selected ontologies: 19759 (reference set for hypergeometric test)

#Genes from Cluster#1: unique uploaded ids 53 -> corresponding genes 52. with 1 (1.89\%) missing -> 52 recognized by ClueGO.

-> To improve the \% of found genes, verify gene identifiers., download new available ClueGO conversion files or add/request additional files.

#Genes with functional annotations in all selected Ontologies from Cluster#1: 52 (100.0\%)

#Genes from all Clusters associated to 48 representative Terms and Pathways (after applying general selection criteria): 34 (65.38\%)

#Genes from all Clusters associated to 13 representative Terms and Pathways (after fusion selection criteria): 26 (50.0\%)

KappaScore Grouping:

Iteration: 0 with 6 groups

Final KappaScore groups = 6

Terms not grouped = 0

Merge redundant groups with >50.0\% overlap

Final group size after merging: 6

#GO All Terms Specific for Cluster #1: 13

#GO Terms: 13

#GO Term Connections: 6

List of missing Genes -> make sure the ids exists or are written properly -> if yes make sure you have their annotation in ClueGO!

V. References

S1. Eickhoff MK, Olsen FJ, Frimodt-Moller M, et al. (2020) Effect of dapagliflozin on cardiac function in people with type 2 diabetes and albuminuria - A double blind randomized placebo-controlled crossover trial. J Diabetes Complications 34(7): 107590. 10.1016/j.jdiacomp.2020.107590

S2. Levey AS, Stevens LA, Schmid CH, et al. (2009) A new equation to estimate glomerular filtration rate. Ann Intern Med 150(9): 604-612. 10.7326/0003-4819-150-9-200905050-00006

S3. Latosinska A, Siwy J, Mischak H, Frantzi M (2019) Peptidomics and proteomics based on CE-MS as a robust tool in clinical application: The past, the present, and the future. Electrophoresis 40(18-19): 2294-2308. 10.1002/elps.201900091

S4. Jantos-Siwy J, Schiffer E, Brand K, et al. (2009) Quantitative urinary proteome analysis for biomarker evaluation in chronic kidney disease. J Proteome Res 8(1): 268-281. 10.1021/pr800401m

S5. Mavrogeorgis E, Mischak H, Latosinska A, Siwy J, Jankowski V, Jankowski J (2021) Reproducibility Evaluation of Urinary Peptide Detection Using CE-MS. Molecules 26(23). 10.3390/molecules26237260

S6. Szklarczyk D, Gable AL, Lyon D, et al. (2019) STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res 47(D1): D607-D613. 10.1093/nar/gky1131

S7. Bindea G, Galon J, Mlecnik B (2013) CluePedia Cytoscape plugin: pathway insights using integrated experimental and in silico data. Bioinformatics 29(5): 661-663. 10.1093/bioinformatics/btt019

S8. Clos-Garcia M, Ahluwalia TS, Winther SA, et al. (2022) Multiomics signatures of type 1 diabetes with and without albuminuria. Front Endocrinol (Lausanne) 13: 1015557. 10.3389/fendo.2022.1015557

S9. Winther SA, Henriksen P, Vogt JK, et al. (2020) Gut microbiota profile and selected plasma metabolites in type 1 diabetes without and with stratification by albuminuria. Diabetologia 63(12): 2713-2724. 10.1007/s00125-020-05260-y

S10. Eder S, Leierer J, Kerschbaum J, et al. (2018) A Prospective Cohort Study in Patients with Type 2 Diabetes Mellitus for Validation of Biomarkers (PROVALID) - Study Design and Baseline Characteristics. Kidney Blood Press Res 43(1): 181-190. 10.1159/000487500

VI. Reporting checklist for randomised trial (CONSORT).

Based on the CONSORT guidelines.

Instructions to authors

Complete this checklist by entering the page numbers from your manuscript where readers will find each of the items listed below.
Your article may not currently address all the items on the checklist. Please modify your text to include the missing information. If you are certain that an item does not apply, please write "n/a" and provide a short explanation.
Upload your completed checklist as an extra file when you submit to a journal.
In your methods section, say that you used the CONSORT reporting guidelines, and cite them as:
Schulz KF, Altman DG, Moher D, for the CONSORT Group. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials

		Reporting Item	Page Number
Title and Abstract			
Title	#1a	Identification as a randomized trial in the title.	1
Abstract	#1b	Structured summary of trial design, methods, results, and conclusions	2
Introduction			
Background and objectives	#2a	Scientific background and explanation of rationale	4
Background and objectives	#2b	Specific objectives or hypothesis	4
Methods			
Trial design	#3a	Description of trial design (such as parallel, factorial) including allocation ratio.	4, 5, Supplementary Material (SM) pages 2-3
Trial design	#3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons	SM pages 2, 3
Participants	#4a	Eligibility criteria for participants	5, SM pages 2, 3
Participants	#4b	Settings and locations where the data were collected	5, SM page 2
Interventions	#5	The experimental and control interventions for each group with sufficient details to allow replication, including how and when they were actually administered	4, 5, SM pages 2, 3
Outcomes	#6a	Completely defined prespecified primary and secondary outcome measures, including how and when they were assessed	4, 5, SM page 3
Outcomes	#6b	Any changes to trial outcomes after the trial commenced, with reasons	n/a
Sample size	#7a	How sample size was determined.	SM pages 3, 4
Sample size	#7b	When applicable, explanation of any interim analyses and stopping guidelines	n/a
Randomization - Sequence generation	#8a	Method used to generate the random allocation sequence.	SM page 3
Randomization - Sequence generation	#8b	Type of randomization; details of any restriction (such as blocking and block size)	SM page 3
Randomization - Allocation concealment mechanism	#9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the	SM page 3

		sequence until interventions were assigned	
Randomization - Implementation	#10	Who generated the allocation sequence, who enrolled participants, and who assigned participants to interventions	SM page 3
Blinding	#11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how.	SM page 3
Blinding	#11b	If relevant, description of the similarity of interventions	n/a
Statistical methods	#12a	Statistical methods used to compare groups for primary and secondary outcomes	6, 7
Statistical methods	#12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses	6, 7, SM pages 2-6
Results			
Participant flow diagram (strongly recommended)	#13a	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome	7, SM pages 7, 8, 11, 12
Participant flow	#13b	For each group, losses, and exclusions after randomization, together with reason	5, SM page 8
Recruitment	#14a	Dates defining the periods of recruitment and follow-up	SM pages 2, 3
Recruitment	#14b	Why the trial ended or was stopped	n/a (It got completed as expected)
Baseline data	#15	A table showing baseline demographic and clinical characteristics for each group	19, SM page 12
Numbers analysed	#16	For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups	20, 21, SM pages 12, 14
Outcomes and estimation	#17a	For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)	SM pages 12, 14
Outcomes and estimation	#17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended	n/a
Ancillary analyses	#18	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory	21, 23-25, SM pages 14
Harms	#19	All important harms or unintended effects in each group (For specific guidance see CONSORT for harms)	SM page 13
Discussion			
Limitations	#20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses	13
Generalisability	#21	Generalisability (external validity, applicability) of the trial findings	8, 21, SM pages 16, 17
Interpretation	#22	Interpretation consistent with results, balancing benefits, and harms, and considering other relevant evidence	9-14

Registration	#23	Registration number and name of trial registry	4, 5, SM page 2
Other information			
Interpretation	#22	Interpretation consistent with results, balancing benefits, and harms, and considering other relevant evidence	9-14
Registration	#23	Registration number and name of trial registry	4, 5, SM page 2
Protocol	#24	Where the full trial protocol can be accessed, if available	14
Funding	#25	Sources of funding and other support (such as supply of drugs), role of funders	19, SM page 2