Electronic Supporting Information

Storage of Serum Peptide Information in Nanoporous Silicon Microparticles

*Jie Tan[a], Xiao Li[a], Guansheng Du[a], Aiwu Pan[b], and Jianmin Wu[a]**

[a]*Institute of Microanalytical System, Department of Chemistry, Zhejiang University, Hangzhou, 310058*

[b]*Hospital of Zhejiang university, Hangzhou, Zhejiang University, Hangzhou, 310058*

Corresponding author: Jianmin Wu

Email address: wjm-st1@zju.edu.cn

1. Method

Preparation of nanoporous Si microparticles (NPSMPs). Porous Si was obtained by anodization of a boron-doped silicon wafer (resistivity 0.5 ~ 1.2 mΩ cm) of [100] crystal orientation in an electrolyte mixture of aqueous hydrofluoric acid (48% by mass, Alatin Corp.), dimethylsulfoxide (DMSO) and distilled water (volume ratio = 2: 4:1). Porosity of samples was controlled by the current density used in the electrochemical etch. The electrochemical cell was rinsed with deionized water after the etching step, and then refilled with a solution containing 3.3% HF (by mass) and ethanol (CAUTION: HF is highly toxic and contact with skin should be avoided). The porous Si film was detached from the silicon substrate by application of a constant current (~18 mA cm^{-2}) for 3 min. The film was placed in a vial containing ethanol. The NPSMPs were obtained after the detached film was fractured in a JY88-IIN (Ningbo Xingyi Ultrasonic Instrument Co. China) ultrasonic bath (for 15 min).

Chemical modification of NPSMPs. Undecylenic acid was grafted to the NPSMPs surface by thermal hydrosilylation[1]. Briefly, the particles were placed in a flask containing a deoxygenated (purge with nitrogen) solution of undecylenic acid (1 mL) dissolved in toluene (10 mL), which was then heated to reflux for 3 h. The unreacted undecylenic acid was removed by rinsing with toluene and ethanol subsequently, and the particles were finally dried in a stream of nitrogen.

Characterization of NPSMPs. The thickness and pore size of the particles was characterized by a field emission scanning electron microscopy (Model Utral 55, CorlzeisD, Germany).

Observation of protein admission/exclusion behavior of porous Si by Fourier transformed reflective interferometric spectroscopy. The ability of a porous Si with a given porosity to admit test proteins was determined by optical interference spectroscopy [2]. A flow cell constructed of Plexiglas was connected to a peristaltic

pump, and protein solutions were introduced to the porous Si chip in a single-pass configuration. The porosity and surface chemistry of porous Si chip is the same as that of microparticle described above. Light from the one arm of optical probe was focused on the surface of the porous Si layer through the Plexiglas cover. Reflected light was collected through the same optics, and the distal end of the second arm of the bifurcated fiber optic cable was input to the CCD spectrometer. Spectra were not corrected for spectral response of the instrument or the lamp. The quantity *2nL*, commonly referred to as the effective optical thickness (EOT), was determined from the Fabry-Pérot relationship:

$$m\lambda = 2nL \qquad (1)$$

where λ is the wavelength of maximum constructive interference for spectral fringe of order *m*, *n* is the index of refraction of the porous layer and its contents, and *L* is the thickness of the porous layer. The value of *2nL* was determined by Fourier transformation of the reflectivity spectrum[3]. The position of the peak in the Fourier transform along the x-axis is equal to the quantity (*2nL*) from eq. (1). The optical interference spectra were recorded at 30 s intervals. A continuous flow (1.0 mL min^{-1}) of pH 7.4 phosphate buffered saline (PBS) solution was pumped through the flow cell until a stable baseline in the optical spectrum was established. The feed solution to the flow cell was then switched to a solution of the protein of interest (1.0 mg mL^{-1}) in PBS. The shift in the *2nL* value, caused by infiltration of protein into the pores of the porous Si film was monitored throughout the entire experimental process. Bovine serum albumin (BSA, (66.2 kDa, Sangon Biotech Co., Ltd.), horseradish peroxidase (HRP, 38.8 kDa, Shanghai Lizhu Dong Feng Biotechnology Co., Ltd), trypsin (24 kDa, Bio Basic Inc.), lysozyme (14.4 kDa, Shanghai Bo'ao Biological Technology Co., Ltd.) and human insulin (5.7 kDa, Wanbang Bio-pharm. Co.) were used as the model proteins to establish the relationship between the cut-off MW of protein and the porosity of porous Si.

Collection of proteins from test solutions or from sera by NPSMPs. Porous Si microparticles (0.3 mg) were added to a vial containing a serum sample diluted with

pH 7.4 PBS buffer, and then shaken at room temperature for 30 min. The suspension was centrifuged at 8000 rpm for 2 min, and then the supernatant was removed from the microparticles. The particles were re-suspended in distilled water and thoroughly washed to remove physic-sorbed species. The particles were isolated by centrifugation (2000 rpm for 1 min) and stored in 25°C or -20 °C for different time as indicated in Table 1.

Direct detection of proteins captured in porous Si microparticles by MALDI-TOF Peptides captured by the NPSMPs were identified by MALDI-TOF MS. The mass spectra were obtained on ultrafleXtreme MALDI-TOF/TOF mass spectrometer (Bruker Daltonics Corp.) equipped with a 337 nm nitrogen laser. The particles were suspended in water and spotted onto the MALDI plate. To each spot was added 1 uL of α-cyano-4-hydroxycinnamic acid (CHCA) dissolved in a mixture of water, acetonitrile and TFA (v/v 50:50:0.1), and the samples were air-dried before insertion into the instrument. Analyses were performed in linear positive-ion mode, using delayed extraction. The acceleration voltage was 20 kV, and the guide wire and grid voltages were maintained at 0.05% and 91.5% of the accelerating voltage, respectively. The delay time was 13800 ns. Samples were evaluated at an m/z range of 300 ~ 10,000 Da. Five hundred laser shots were averaged for each mass spectrum. The raw spectra were processed with FlexAnalysis software (Bruker Daltonics Corp.)

Identification of peptides captured in porous Si microparticles by nano-LC-ESI MS/MS All LC separations were performed on a MS4-APIQSTAR NANOLC-Q-TOF system (ABI Applied Biosystems. Michrom Bioresources). The analysis was performed under the following conditions: the peptide samples (10 μL) were enriched by NPSMPs and eluted with 20 μL 50% ACN after washed by water. The peptides were automatically loaded across a Paradigm Platinum Peptide Nanotrap (Michrom Bioresources, Inc.) precolumn (0.15 × 50 mm) for sample concentrating at a flow rate of 1 μL/min in buffer A (98% pure water/2% ACN, containing 0.1% formic acid). The in-line analytical capillary column (75 μm × 12 cm) was packed

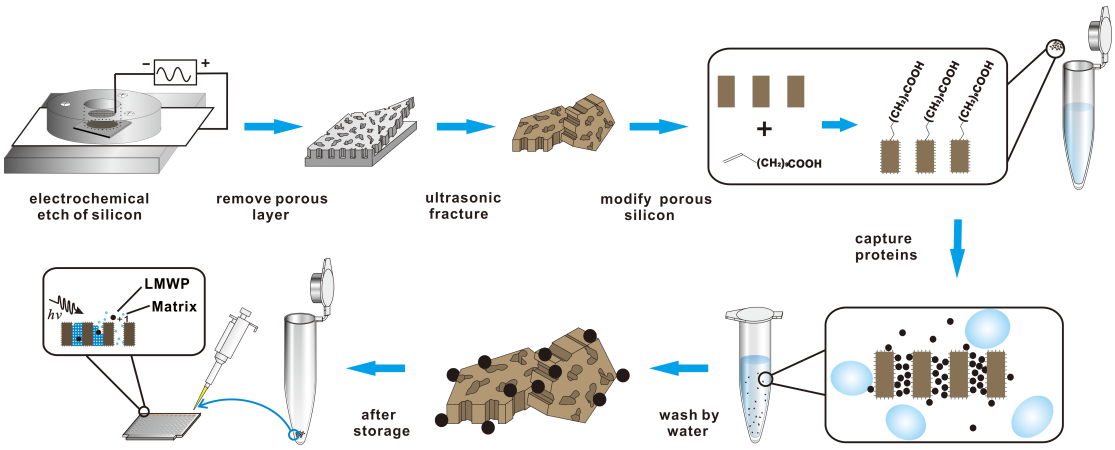
with C18 stationary phase (5 μm, 200 Å Magic C18AG, Michrom Bioresources, Inc.) and Picofrit capillary tubing (New Objective, Cambridge, MA). Peptides were eluted using a linear gradient of 5–30% buffer B (2% pure water/98% ACN, containing 0.1% formic acid) over 8 min followed by a linear gradient of 30–50% buffer B for 10 min, then buffer B was ramped to 65% for 15 min and increased to 90% for 5 min before declining to 5% across the column. The column was connected via a fused silica capillary to a low-volume tee (Upchurch Scientific) where a high voltage (2600 V) was applied and the column tip was positioned approximately 0.5 cm from the Z-spray inlet of an Ultima API hybrid Q-TOF tandem mass spectrometer (Micromass, Manchester, UK). Positive ions were generated by electrospray, and the Q-TOF operated in data-dependent acquisition mode. A TOF MS survey scan was acquired (m/z 350-1700, 1 s), and the 3 largest multiply charged ions were sequentially selected by Q1 for MS/MS analysis. Argon was used as collision gas and optimum collision energy was chosen (based on charge state and mass). Tandem mass spectra were accumulated for up to 8 s (m/z 50-2000). LC-MS/MS data were processed by Analysis software (ABI applied biosystem Co.). The mass spectra of each sample were obtained by selecting the full range of total ion chromatogram (TIC).

Criteria for database Search The acquired MS/MS spectra were searched against the Swiss-Prot database with the ProteinPilot^{TM} software 2.0.1 (Applied Biosystems Co.), in which no enzyme was selected in digestion section. The human protein database (Homo sapiens taxonomy) was used in this study. The detected protein threshold was > 0.10 (20%), the competitor error margin was set as 2.00. All Protein Pilot Paragon analyses were conducted using simultaneous analysis against a reversed SwissProt database and automated filtering and retaining of peptides scoring with q-values (false discovery rate filter) ≤ 0.10.

Disease classification using cluster analysis The raw spectra were processed using Bruker Daltonics flexAnalysis software and the data were exported to MATLAB for pretreatment. All spectra were aligned using an m/z tolerance of 0.15% and peak

intensity was normalized. The baseline was corrected and the negative values were removed prior to analysis. Hierarchical clustering was performed using SPSS and visualized by MATLAB. The MALDI data (m/z peak intensities) were normalized and median centered. Pearson correlation was used to calculate the distance between the samples, and complete linkage clustering was performed.

2. Supplementary Figure



Scheme S1 Schematic workflow of NPSMPs preparation, surface modification and application in the storage of serum peptide information. A layer of porous Si was prepared by electrochemical etching a crystalline silicon wafer. The porous layer was then removed and fractured into microparticles in an ultrasonic bath. The particles were chemically modified to impart affinity for peptides and to stabilize the porous nanostructure. The microparticles were incubated with serum sample, followed by harvesting and storing at different conditions. The microparticles with captured peptides can be directly spotted on a MALDI plate and the fingerprint of serum peptide can be obtained with a MALDI-TOF mass spectrometer.

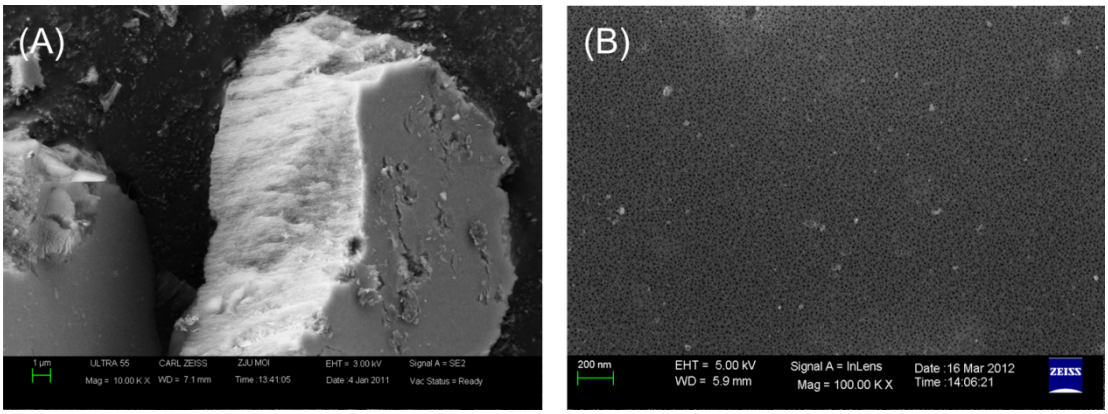


Fig. S1 (A) Cross-sectional SEM image of porous Si microparticles modified by hydrosilylation with undecylenic acid. (B) Plan-view of scanning electron microscope (SEM) image of carboxyl-terminated porous Si microparticles. The pore diameter and thickness of the porous Si is respectively ~9 nm and 8μm, which was measured by SEM. The porosity can also be estimated by calculating the percentage of pore space occupied in the total volume of the particles.

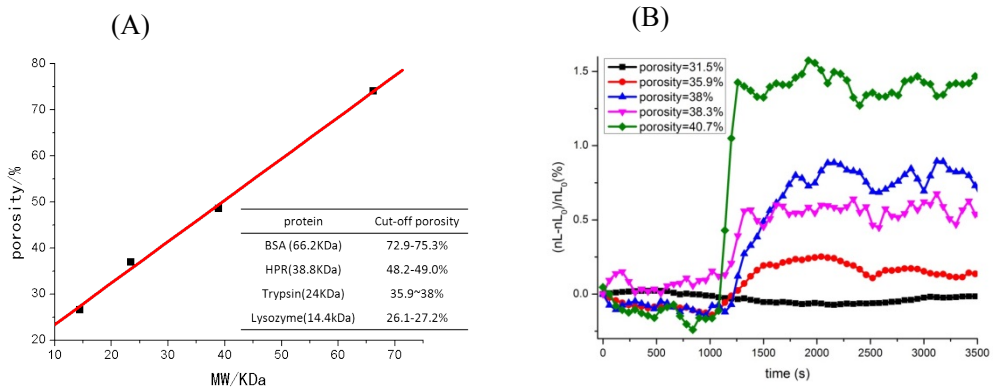


Fig. S2 (A) The relationship between protein molecular weight and the critical porosity of the porous Si sample that is able to admit the protein. The critical porosity represents the minimum or cut-off value, below which the indicated protein will not enter the porous Si matrix. Protein admission or exclusion by the porous Si can be monitored by Fourier transformed reflective interferometric spectroscopy (FT-RIS). (B) Determination of critical porosity for excluding trypsin molecules by FT-RIS. The sensorgram indicated that porous Si with porosity smaller than 31.5% can fully exclude trypsin molecules and probably other protease. Consequently, the captured peptides can be protected.

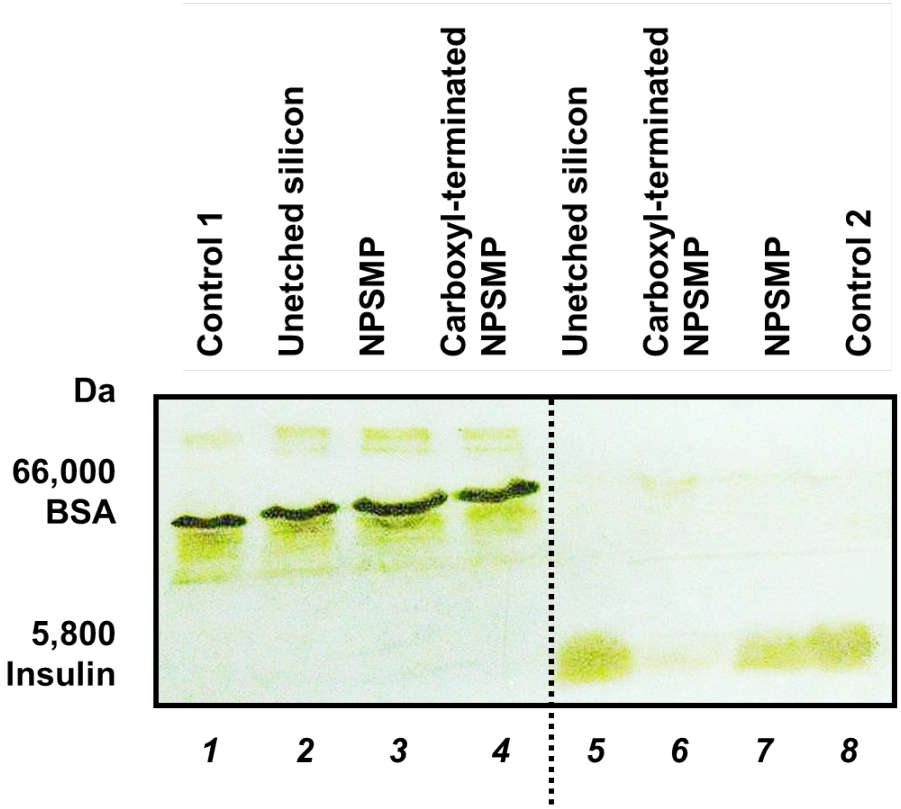


Fig. S3 SDS-PAGE analysis of proteins remaining in test solutions treated with different types of particles. Lane 1 to 4 display the SDS-PAGE results for the solution containing BSA with initial concentration of 1.0 mg·mL^{-1} (1) solution prior to particle introduction (Control 1); (2) supernatant from unetched silicon microparticles; (3) supernatant from NPSMPs without surface modification (native Si-H surface); (4) supernatant from carboxyl-terminated NPSMPs. The result indicated that BSA (MW=66 kDa) can be excluded by each types of microparticles. Lane 5 to 8 display the SDS-PAGE results for the solution containing insulin with initial concentration of 20 U·mL^{-1} (5) supernatant from unetched silicon microparticles; (6) supernatant from carboxyl-terminated NPSMPs; (7) supernatant from NPSMPs without surface modification (native Si-H surface); (8) solution prior to particle introduction (Control 2).The result clearly indicated that insulin with molecular weight of 5.8 kDa can be captured by the carboxyl-terminated NPSMPs.

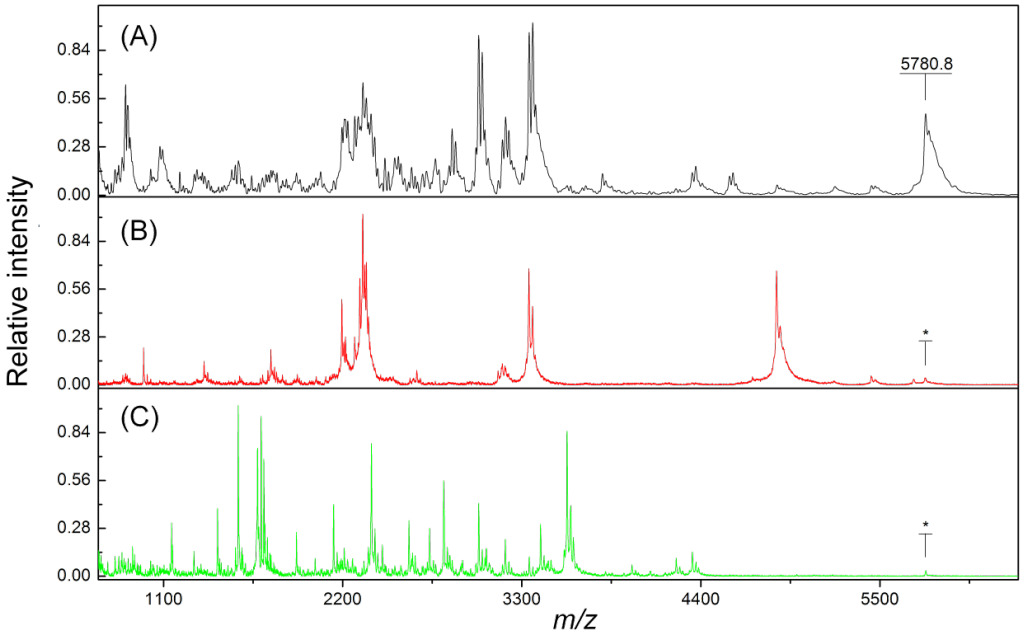


Fig. S4 MALDI-TOF mass spectra of insulin (MW = 5.8 kDa) in the presence of trypsin (MW = 24.0 kDa) and NPSMPs with different porosity after 2 h incubation: (A) mass spectrum of mixed sample (insulin + trypsin) in the presence of NPSMPs with a porosity of 26% (pore diameter ≈ 9 nm); (B) mass spectrum of mixed sample in the presence of NPSMPs with porosity of 50% (pore diameter ≈ 14 nm); (C) mass spectrum of mixed sample in the absence of NPSMPs (Control). The results show that the insulin peak can be only found in the sample protected by NPSMPs with porosity of 26%, which is below the critical porosity for admission of trypsin. Larger pore size allows the admission of protease into the pore channel, causing the degradation of insulin.

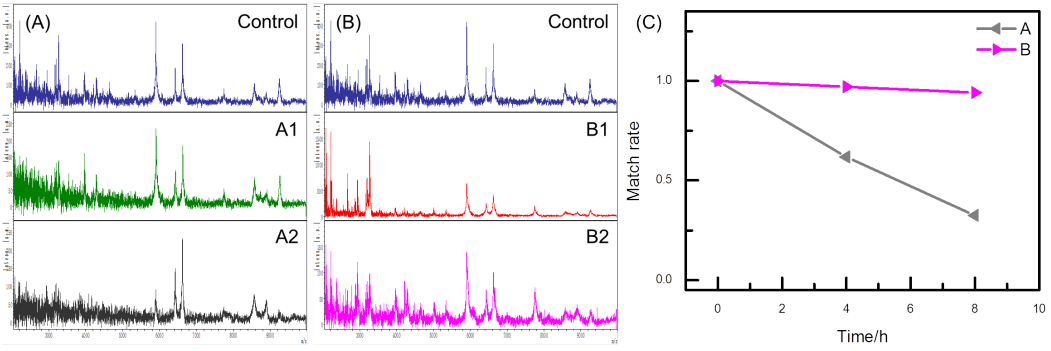


Fig. S5 Peptides fingerprint of human serum kept at 4 °C. (A) The peptide profile of serum sample without pretreatment by NPSMPs. The samples were stored for 4 hours (A1) and 8 hours (A2), respectively; (B) The peptide profile of serum sample pre-treated by NPSMPs. The samp C stored for 4 hours (B1) and 8 hours (B2), respectively. Control sample in both (A) and (B) was processed and analyzed immediately after it was enriched with NPSMPs; (C) Match rate of sample without pre-treatment(Gray trace), and pre-treated sample (Pink trace) during the short-term storage.

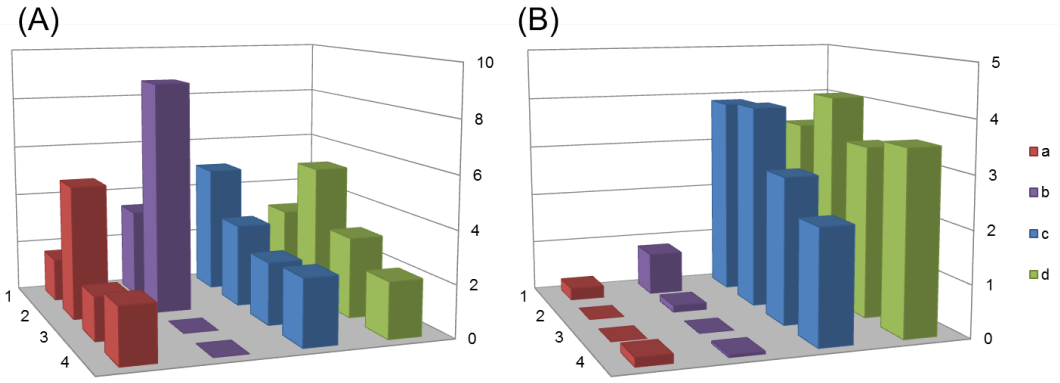


Fig. S6 The change of mass spectroscopic intensity in different storage conditions. (A) The change of mass spectroscopic intensity in the MW range from 2 to 6 kDa. (B) The change of mass spectroscopic intensity in the MW range from 6 to 10 kDa. Each bar represents the sum of normalized peak intensity. The red bar, purple bar, blue bar and green bar represent the sample categories of group a, b, c and d, respectively. Each sample group was stored at conditions listed in Table 1.

Table S1 Peptides (Mr. > 1000) identified (Conf > 1.3) by LC-MS/MS*

Prec MW	Best Sequence	Modifications
1205.597	EGDFLAEGGGVR	
1205.597	EGDFLAEGGGVR	
1205.606	EGDFLAEGGGVR	
1205.608	EGDFLAEGGGVR	
1262.634	GEGDFLAEGGGVR	
1308.554	DSGEGDFLAEGGGV	
1324.534	QGVNDNEEGFFS	Gln->pyro-Glu@N-term
1330.545	DSGEGDFLAEGGGV	Cation:Na(E)@10
1330.583	DSGEGDFLAEGGGV	Cation:Na(E)@10
1349.674	SGEGDFLAEGGGVR	
1384.841	TSIGHVVQLLYR	
1395.584	QGVNDNEEGFFSA	Gln->pyro-Glu@N-term
1417.559	QGVNDNEEGFFSA	Gln->pyro-Glu@N-term, Cation:Na(E)@8
1455.846	LVLLALEDGSALSR	
1464.671	DSGEGDFLAEGGGVR	
1464.694	DSGEGDFLAEGGGVR	
1464.701	DSGEGDFLAEGGGVR	
1464.706	DSGEGDFLAEGGGVR	
1464.707	DSGEGDFLAEGGGVR	
1481.572	ADSGEGDFLAEGGGV	Phospho(S)@3, Cation:Na(E)@11
1486.67	DSGEGDFLAEGGGVR	Cation:Na(E)@4
1517.68	ADSGEGDFLAEGGGVR	Dehydrated(S)@3
1535.754	ADSGEGDFLAEGGGVR	
1544.629	DSGEGDFLAEGGGVR	Phospho(S)@2
1544.679	DSGEGDFLAEGGGVR	Phospho(S)@2
1615.675	ADSGEGDFLAEGGGVR	Phospho(S)@3
1615.685	ADSGEGDFLAEGGGVR	Phospho(S)@3
1615.714	ADSGEGDFLAEGGGVR	Phospho(S)@3
1615.714	ADSGEGDFLAEGGGVR	Phospho(S)@3
1615.72	ADSGEGDFLAEGGGVR	Phospho(S)@3
1933.148	SKITHRIHWESASLLR	
2020.154	SSKITHRIHWESASLLR	
2020.187	SSKITHRIHWESASLLR	
2157.128	SFTLASSETGVGAPISGPGIPGR	
2200.217	LQTPASFAQSVQELTIALQR	
2200.251	LQTPASFAQSVQELTIALQR	
2200.279	LQTPASFAQSVQELTIALQR	

2426.078	VVGTAWTADSGEGDFLAEGGGVR	Oxidation(W)@6, Phospho(T)@7, Phospho(S)@10
2745.5	TVTELILQHQNPQQLSSNLWAAVR	
2930.435	SSSYSKQFTSSTSYNRGDSTFESKSY	
3189.525	SSSYSKQFTSSTSYNRGDSTFESKSYKM	
3189.527	SSSYSKQFTSSTSYNRGDSTFESKSYKM	
3260.58	SSSYSKQFTSSTSYNRGDSTFESKSYKMA	
3260.6	SSSYSKQFTSSTSYNRGDSTFESKSYKMA	
3260.603	SSSYSKQFTSSTSYNRGDSTFESKSYKMA	
3439.648	DFLAEGGGV	
3658.963	LRPHLELLANIDPSPDAPPPTWEQLENGLVAVR	
3953.142	QAGAAGSRMNFRPGVLSSRQLGLPGPPDVPDHAAYHPF	Gln->pyro-Glu@N-term, Deamidated(N)@10
3156.638	NVHSGSTFFKYYLQGAKIPKPEASFSPR	Deamidated(Q)@14
3954.17	QAGAAGSRMNFRPGVLSSRQLGLPGPPDVPDHAAYHPF	Gln->pyro-Glu@N-term, Deamidated(N)@10
1459.603	ADSGEGDFLAEGGGV	Phospho(S)@3
1707.966	TVVHGLVDYIQNHSK	
1439.539	EDQEADGEEDEGT	Oxidation(D)@2, Deamidated(Q)@3
1019.533	DFLAEGGGVR	
1259.496	QGVNDNEEGFF	Gln->pyro-Glu@N-term, Cation:Na(E)@8
1864.051	SSKITHRIHWESASLL	
1349.654	SGEGDFLAEGGGVR	
1566.788	DSGWRDLFSRSEP	Oxidation(P)@13
1076.555	GDFLAEGGGVR	
2222.263	LQTPASFAQSVQELTIALQR	Cation:Na(E)@13
1089.626	LVMGIPTFGR	
1933.134	SKITHRIHWESASLLR	
1637.696	TADQARVGPADDGPAPS	Methyl(T)@1
1346.537	PLDREEQPEFS	Deamidated(Q)@7
1816.965	THGFDGLDLAWLYPGR	
1777.061	SKITHRIHWESASLL	

1619.874	SSKITHRIHWESAS	Dehydrated(S)@12
1864.071	GTVTSGKRKVYLLSLLL	Oxidation(K)@7, Deamidated(R)@8
1831.012	QLQELALEERQTIDQ	Deamidated(Q)@1, Deamidated(Q)@3, Oxidation(R)@10
1864.053	TGRSLEHRPRGNTASLI	
1532.844	GAVVGVPARWHRGR	Oxidation(H)@11
1135.701	PTPDTGPQNGH	Oxidation(H)@11
2222.252	IQTALSSKDGRLPRTYRLF	Deamidated(R)@14
1196.758	GSPGGPGVSGGSPAGG	
2713.136	SPIRTMTDILSRGPKSMISLAGGLPN	Deamidated(R)@4, Deamidated(R)@12
1114.598	IKENQKDLQ	
1262.908	MAKPIGLCMPN	Oxidation(K)@3, Oxidation(P)@4, Carbamidomethyl(C)@8
3260.641	PLDASPRRPPGPTTSPASTSLSSPGQRDDLIA	Deamidated(R)@8, Oxidation(P)@24
1020.654	MEDVNSNVN	
1436.909	VTVPLPASQLSLPN	Deamidated(Q)@9, Deamidated(N)@14
1604.906	THRIHWESASLLR	
1866.064	CNMGNLICDAMINNNL	Carbamidomethyl(C)@1 , Carbamidomethyl(C)@8
1973.022	EQLAVGGSLVQPAVAPSSGGAP	Glu->pyro-Glu@N-term
2959.602	TGAASGIMSLAGLVLAPFTAGTSLALTAAGVGL	
1020.652	SSQNRKSGSA	
1117.263	AGTTVKVWDAA	
1158.538	AGKGGGGDGGGQGGK	
1240.652	AAAGGGAKSRPLAN	Deamidated(N)@14
1275.466	FCEQQSNDTF	Carbamidomethyl(C)@2 , Deamidated(N)@7
1297.371	GRSAPGAASAAAALK	
1414.891	KNAQPSPARAPAE	Phospho(S)@6
1431.776	KQSALMVNGVLKQ	Oxidation(K)@1, Deamidated(Q)@2
1530.743	RGLGGPAPPEPDSGPQ	
1655.903	QRPGLQVILMSATLN	Oxidation(P)@3

1790.061	SLGSDGDCAANSPAAGGGAR	Carbamidomethyl(C)@8
1831.167	ATGAASQSPLPQYVTVKGG	
2012.229	PQPTIPIVGIVAGLVVLGAVV	Deamidated(Q)@2
2077.201	SALVGRTNGLTKPAALAAAPAK	
2108.893	FAKETSLLAVPGALSPLAIPN	
2868.023	SSVLGLSYLQGGGAGSASGGASGGSSGGAASGAGPG	Oxidation(P)@35
5134.257	K.QFTAVKNDYEMTFNNETSVMPCEDDHHLPTVQFDFTGIDDLENK.S	
6435.244	SVAWGWSRINHHTWHIVGLFAIGLLLAMLRGNHIGHVENWYLIGFAALVFFVLIRD	
6628.533	AVPRVDDEPRAQLGALLARYIQQARKAPSGRMSVIKNLQSLDPSHRISDRDYMGWMDF	
8914.996	NPMQSIMLANLRPSPTTTTTTTTPAAAPTTTAATIALQANDKLQALTPPMDVTPPKSPAKSQQSCAEPEKEHDLMSNSSEDMKY	
9419.461	MTFIRPWNRYNRGETAGFDPATAAGLIGVHAVPYRPAEGAPAAIPAADPAPEPAEPDPSFETAMAALETPAETAPAAAAPGTADLPVQGRRK	
9717.027	PAPRLPQDPCGPDCAPPAPGLPRGPCGPDCAPAAPGLPPDPCGPDCAPPAPGLPQDPCGPDCAPPAPGLPRGPCGPDCAPAAPGLPPDPCGSNCAP	

* The sequences marked with gray background represent the peptides matched with the peaks found in MALDI-TOF MS spectra, (MS tolerance was set as 0.15%). The sequences highlighted with red represent the lost peptide (within HMW range) in the non-protected sample during storage.

Table. S2 Number of peptides identified by different search criteria*

Unused (Conf) Cutoff	**Number of peptides identified based on LC MS/MS**	**Number of peptides match with MALDI spectra**
>2.0 (99)	84	40
>1.3 (95)	105	64
Cutoff Applied: >0.10 (20%)	1259	340

*Match MS tolerance was set as 0.15%.

Reference

1. M. P. Schwartz, F. Cunin, R. W. Cheung, and M. J. Sailor, phys. stat. sol. (a), 2005, **8**, 1380.

2. A. Janshoff, K. P. S. Dancil, C. Steinem, D. P. Greiner, V. S. Y. Lin, C. Gurtner, K. Motesharei, M. J. Sailor, M. R. Ghadiri, J. Am. Chem. Soc. 1998, **120** , 12108.

3. J. Wu, M. J. Sailor, Adv. Funct. Mater. 2009, **19,** 733