# PreON automated sample preparation


**Authors:** Sebastian Steiner[1], Sebastian Johansson[2], Doris Jansen[2], Jonas Albinus[1], Sandra Goetze[1], Harald Quintel[3], Konstantin Lutze[3], Bernd Wollscheid[1], Fabian Hosp[2], Nils Kulak[2]

[1] ETH Zürich, Institute of Molecular Systems Biology, Zurich, Switzerland; [2] PreOmics GmbH, Planegg/Martinsried, Germany; [3] HSE AG, Hombrechtikon, Switzerland

## Introduction

Sample preparation is an often overlooked yet very important component of the overall mass spectrometry-based proteomics workflow and remains to be a limiting factor for reproducible high-throughput analyses. One method to increase the throughput is sample multiplexing using chemical labeling approaches such as iTRAQ or TMT. Here, we present an automated end-to-end solution for standardized sample processing of cells and liquid biopsies, including cell lysis, digestion, TMT labeling and peptide cleanup with only 5 minutes hands-on time. In order to minimize sample loss, improve reproducibility and efficiency, we sought to completely automate both label-free and chemical labeling sample processing from cell lysis to ready-to-measure peptides. To facilitate this, we aimed to combine both the iST and iST-NHS technologies [1] with a newly developed automation platform called PreON (HSE AG).

The workflow described here integrates robotic handling with either label-free or chemical labeling of peptides including cell lysis, protein denaturation, reduction, alkylation, digestion and the peptide cleanup. This solution increases the sample reproducibility, minimizes variability as well as sample loss and dramatically reduces hands-on time compared to manual processing. The PreON platform is flexible to work with 4-16 samples and features a built-in centrifuge, ultrasound liquid detection, deck load check and a drag-and-drop method selection for easy, fast and a convenient menu-driven guidance.

**Keywords**

Sample preparation; automation; proteomics; label-free; TMT labeling; mass spectrometry (MS); serum; mammalian tissue; yeast; HeLa; Ramos B cells

**Key takeaways**

Automated processing of various sample types employing the iST & iST-NHS technology for highly reproducible and sensitive proteomic analyses.


**Figure 1 | PreON platform**

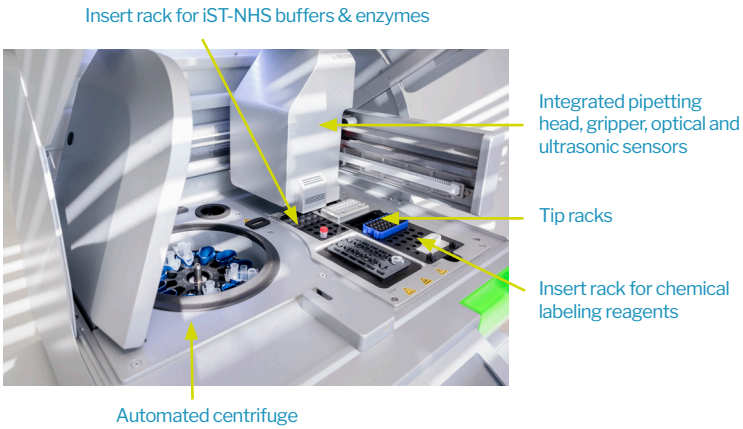


**Figure 2 | PreON workdeck layout**

## Material and Methods

**Sample types and input amounts used in this application note:**

1 $OD_{600}$ yeast pellets (*S.cerevisiae*); 6E5 HeLa cell pellet (*H.sapiens*); 1E4-5E5 Ramos B cells (*H.sapiens*); 1-100 µg standardized serum (*H.sapiens*; Sigma-Aldrich #H4522); 500 µg freeze-dried liver powder (*H.sapiens*; NIST Candidate RM 8461 human liver for proteomics).

Sample preparation was done on the PreON platform or by a manual operator, employing either iST or iST-NHS kits according to protocol. For Ramos B cells and tissue preparation, lysis was performed off-board by adding 50-100 µL of LYSE (PreOmics GmbH) to the sample followed by ultrasonication using the Diagenode Bioruptor® Pico (10 cycles, 30 sec ON/OFF). After the sonication treatment, the lysate was then transferred to the PreON before continuing with a modified PreON protocol without the lysis step.

**Samples were measured using different LC-MS & data analysis setups:**

1. Serum label-free preparation: EASY-nLC 1000 coupled to a QExactive Plus (Thermo Fisher Scientific) with modified settings for data dependent acquisition: MS1 resolution 70,000; MS1 AGC target 1E6; max. injection time 80 ms; MS2 resolution 17,500; MS2 AGS target 5E4; max injection time 60 ms; topN 15; Isolation window 1.4 m/z; NCE 25. Peptides were eluted with a linear 90 min gradient. Column was at ambient temperature of 20°C. Raw files were analyzed using ProteomeDiscoverer v2.1.0.81 (Thermo Fisher Scientific) and the Trans-Proteomic Pipeline. The false discovery rate was set to 0.01 for both proteins and peptides with a minimum length of six amino acids, and was determined by searching a decoy database. Enzyme specificity was set as semi-enzymatic C-terminal to arginine and lysine, using trypsin as the protease, and a maximum of two missed cleavages were allowed in the database search. Label-free quantitation (LFQ) was performed with Progenesis QI (Nonlinear Dynamics).

2. Ramos B cell label-free preparation: EASY-nLC 1200 coupled to a Q Exactive HF-X (Thermo Fisher Scientific) set up for data independent acquisition (DIA): A) For DDA library generation the duplicates of the highest cell input (500k cells) were pooled and remeasure six consecutive times. The data acquisition parameters were set to MS1 resolution 60,000; MS1 AGC target 3E6; max. injection time 15 ms; MS2 resolution 15,000; MS2 AGS target 1E4; max injection time 110 ms; topN 10; isolation window 1.2 m/z; NCE 28. Column was at ambient temperature of 20°C. Raw files were analyzed using ProteomeDiscoverer v2.1.0.81 (Thermo Fisher Scientific) with MS Amanda and Sequest HT as search engine. The false discovery rate was set to 0.01 for both proteins and peptides with a minimum length of six amino acids, and was determined by searching a decoy database. Enzyme specificity was set as C-terminal to arginine and lysine, using trypsin as the protease, and a maximum of two missed cleavages were allowed in the database search. B) For DIA data acquisition parameters were set to MS1 resolution 120,000; MS1 AGC target 3E6; max. injection time 50 ms; scan range 400 to 1210 m/z, DIA scan resolution 30,000; AGS target 1E6; max injection time auto; loop count 18, MSX count 1; 54 isolation windows of 15 m/z; isolation offset 0 m/z, NCE 28. Column was at ambient temperature of 20°C. Raw files were analyzed using Spectronaut Professional+ v12.0.20491.22 (Biognosys). False discovery rate was set to 0.01 and data was normalized with local normalization.

3. Human liver (NIST Candidate RM 8461) tissue label-free preparation: EASY-nLC 1200 coupled to a timsTOF Pro (Bruker Daltonik) with standard settings. Peptides were eluted with a linear 60 min gradient. Column temperature was kept at 60°C. Raw files were analyzed using MaxQuant v.1.6.6.0 [2] with standard settings for ion mobility data. The false discovery rate was set to 0.01 for both proteins and peptides with a minimum length of seven amino acids, and was determined by searching a reverse database. Enzyme specificity was set as C-terminal to arginine and lysine, using trypsin as the protease, and a maximum of two missed cleavages were allowed in the database search. Label-free quantitation (LFQ) was performed with a minimum ratio count of 1. Statistical analysis was performed using Perseus [3].

4. Yeast & HeLa TMT 11-plex preparation: EASY-nLC 1200 coupled to a Q Exactive HF-X (Thermo Fisher Scientific) with modified settings: MS1 resolution 120,000; MS1 AGC target 3E6; max. injection time 50 ms; MS2 resolution 60,000; MS2 AGC target 1E5; max. injection time 120 ms; isolation width 0.8 Th; NCE 32; first fixed mass 100; underfill ratio 2%. Peptides were eluted with a linear 100 min gradient. Column temperature was kept at 60°C. Raw files were analyzed using MaxQuant v.1.6.0.16 [2], the false discovery rate was set to 0.01 for both proteins and peptides with a minimum length of seven amino acids, and was determined by searching a reverse database. Enzyme specificity was set as C-terminal to arginine and lysine, using trypsin as the protease, and a maximum of two missed cleavages were allowed in the database search. TMT 11-plex was set as isobaric labels. For determination of TMT labeling efficiency, the TMT labels were set as variable modifications. For all TMT experiments, a specific cysteine modification ($C_6H_{11}N_O$, +113.084 Da) was considered as fixed modification. Statistical analysis was performed using Perseus [3].

## Results

To assess the automated sample processing performance of our iST Kits, we compared several parameters for either manual or automated kit operation on the PreON instrument.

First, we employed a triplicate dilution series of standardized human serum from 100 µg down to 1 µg protein input. While the manual operation resulted in almost the same number of protein identifications over the whole concentration range, numbers of identified proteins decreased slightly for input amounts of 10 µg and less when processed on the PreON (Figure 3). The reduced number of identifications for low input samples was related to the prototype software, which has been updated in the most recent version of the PreON to optimize the transfer of sample volumes.

Second, we compared the technical variability of these triplicate dilution series experiments and found that the PreON achieves lower coefficients of variation (CVs) compared to manual operation and thus a more precise handling of the samples (Figure 4).

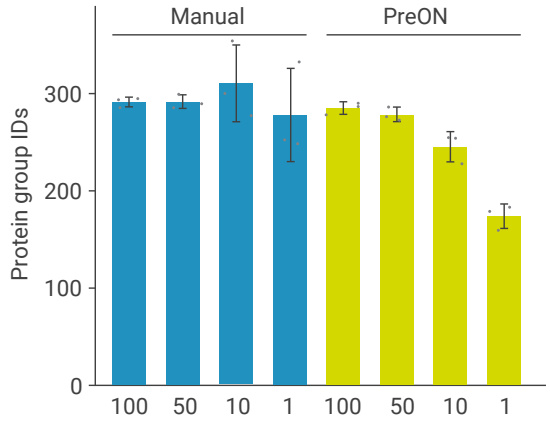


**Figure 3 | Dilution series of human serum. Manual iST sample preparation vs. operation on PreON.**

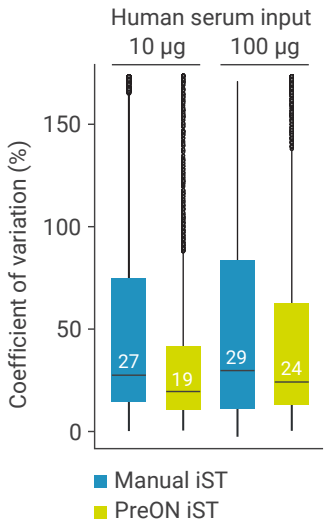


**Figure 4 | Technical variability of triplicates from Fig. 1 for 10 & 100 µg serum input. Manual iST sample preparation vs. operation on PreON.**

Third, we evaluated the quantitative reproducibility for both handling methods and found that the PreON resulted in higher reproducibility than manual operation at low input amounts (Figures 5 and 6).

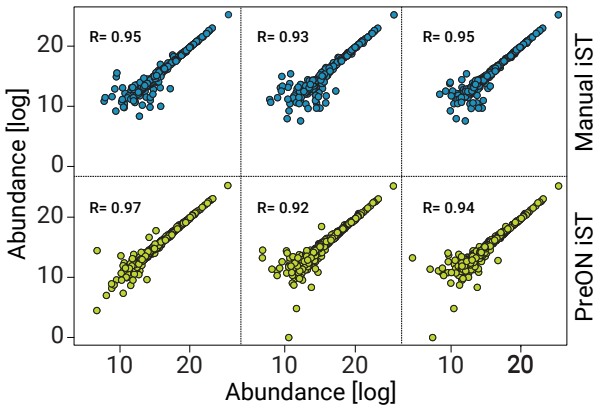


**Figure 5 | Quantitative reproducibility of triplicates from Fig. 3 for 100 µg serum input. Manual iST sample preparation (green) vs. operation on PreON (blue).**

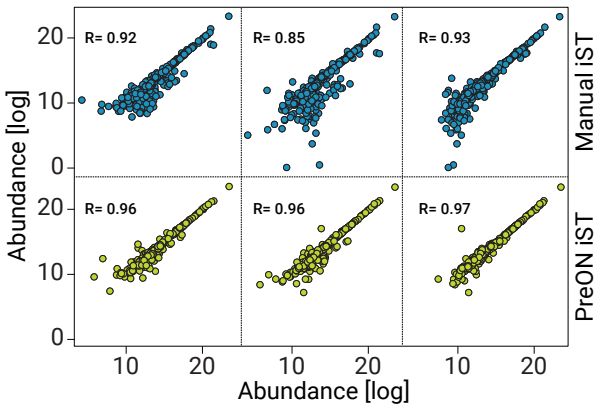


**Figure 6 | Quantitative reproducibility of triplicates from Fig. 3 for 10 µg serum input. Manual iST sample preparation (green) vs. operation on PreON (blue).**

Fourth, we assessed the intra-day, inter-day and inter-instrument reproducibilities for sample processing on the PreON and observed excellent Pearson correlation across all conditions (Figure 7).

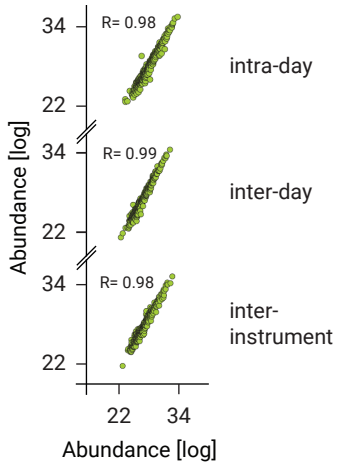


**Figure 7 | Reproducibility of PreON runs, each run done with yeast pellets, 100 µg protein input, biological quadruplicates.**

Intra-day reproducibility of two sample processing runs on the same day and same PreON instrument. Inter-day reproducibility of two sample processing runs on two consecutive days on the same PreON instrument. Inter-instrument reproducibility of two sample processing runs on the same day on two distinct PreON instruments. Pearson correlation displayed in red.

Next, we asked how the PreON performs for cultured mammalian cells and assessed the input sensitivity. To this end, we performed a dilution experiment with Ramos B cells (human Burkitt's lymphoma) in duplicates, starting from 500,000 cells (equal to roughly 100 µg protein input) down to 10,000 cells (equal to roughly 2 µg protein input). We found that the number of protein identifications remained stable at more than 3,000 proteins over a range of 500,000 to 100,000 intact cells (Figure 8). For 50,000 Ramos B cells (~10 µg protein), the identifications decreased slightly and, for 10,000 cells, decreased to about 2,000 proteins.

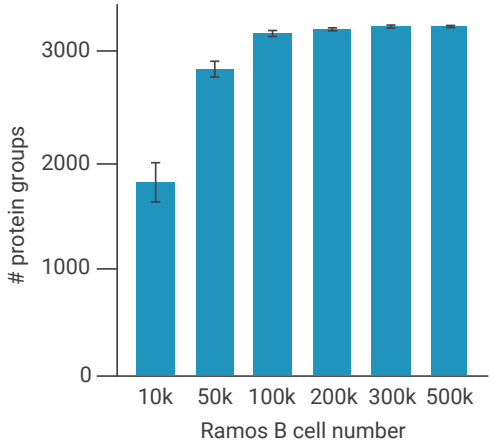


**Figure 8 | Dilution series for 10k-500k human Ramos B cells processed on the PreON. 500,000 Ramos B cells equal ~100 µg protein input; biological duplicates.**

We evaluated further input material and processed a newly developed human tissue standard on the PreON, the NIST Candidate RM 8461 human liver for proteomics. Since this sample type is more rigid and difficult to lyse, we performed the lysis step outside of the PreON by adding LYSE buffer to the sample, boiling it for 10 min at 95°C and then subjected it to repeated cycles of ultrasonication for a total 10 min before continuing with digestion and peptide cleanup on the PreON platform. We observed high reproducibility across quadruplicates (average Pearson correlation R>0.91; Figure 9), indicating that the PreON is capable of handling even more complex sample matrices when combined with off-board lysis.

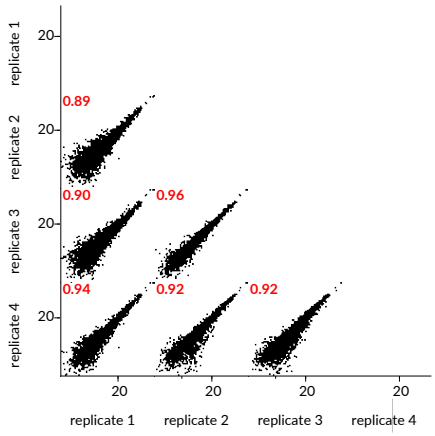


**Figure 9 | Reproducibility of standardized human liver lysates processed on the PreON; Pearson correlation displayed in blue; 50 µg protein input.**

All experiments described before were performed with label-free quantification. To increase throughput of sample preparation, we aimed to implement sample multiplexing using chemical labeling on the PreON. For this purpose, a second tube decking station (Figure 2) is utilized to store tandem mass tags vials (TMT), acetonitrile for TMT resolubilization and hydroxylamine for quenching the TMT reaction.

We employed 11-plex TMT labeling with yeast and HeLa cells and measured a labeling efficiency of >98% for all 11 TMT channels (Figure 10). Thus, the PreON allows fully automated sample processing of cultured human cells, including cell lysis, protein denaturation, reduction, alkylation, digestion, resuspension of lyophilized TMT reagents, 11-plex TMT labeling and quenching, as well as the peptide cleanup.

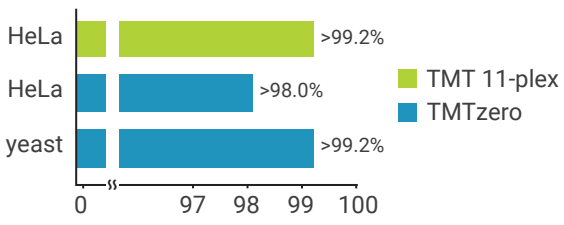


**Figure 10 | TMT labeling efficiency for yeast and HeLa samples, 50 µg protein input. Lysis, digestion, TMTzero or TMT 11-plex labeling and peptide cleanup all performed on the PreON instrument.**

## Conclusions

Automated sample preparation on the PreON drastically reduces hands-on time down to five minutes for both label-free and TMT applications (Figure 11). The workflow here employed enables sample processing and TMT labeling of up to 11 samples in parallel in an automated fashion and in less than 4 hours, scaling from 1-100 µg of protein input material. Automation of our manual iST workflows on the PreON delivers constant results as demonstrated by the excellent Pearson correlation for intra-day, inter-day and inter-instrument assessments.

Preinstalled menu-driven workflows also enable users without any special training or prior knowledge to obtain the same results as an experienced protein analysis user. The intuitive user interface also allows flexible applications for chemical labeling experiments such as 2-plex, 6-plex, 10-plex, 11-plex or 16-plex sample processing. Currently, the PreON provides a third iST workflow, the iST-BCT, which can be easily selected by the provided software interface on the PreON touchsceen.

Furthermore, every step of the PreON workflow is entirely traceable. This is particularly valuable in regulated laboratory environments. Kits from PreOmics or chemical labeling reagents containing a barcode can be scanned by a provided barcode scanner. All information is stored on the PreON and can be downloaded or transfered to a laboratory infromation management system.

We foresee the PreON evolving as the flexible platform for reproducible and standardized sample processing. Label-free and chemical labeling workflows are the first in a pipeline of applications that allow walk-away automation for mass spectrometry-based protein analysis.

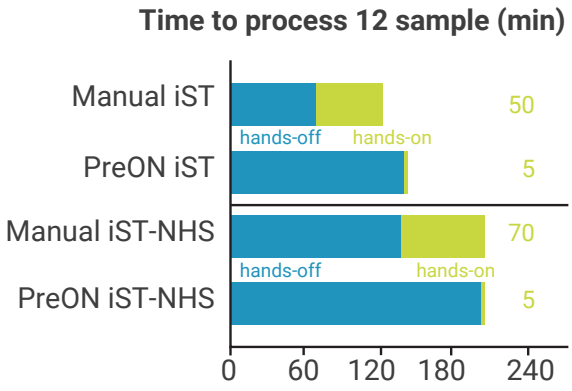


**Figure 11 | Comparison of hands-on and hands-off times for a manual operator vs. operation on the PreON instrument. Top chart for label-free applications, bottom chart for chemical labeling.**

## Products

| Product | Manufacturer | Product Code |
|---|---|---|
| PreON Instrument | HSE-AG / PreOmics GmbH | P.O.00087 |
| iST PreON 8x | PreOmics GmbH | P.O.00077 |
| iST PreON 96x | PreOmics GmbH | P.O.00079 |
| iST -NHS PreON 12x | PreOmics GmbH | P.O.00078 |
| iST-NHS PreON 96x | PreOmics GmbH | P.O.00080 |
| PreON 200 µL pipette tips | PreOmics GmbH | P.O.00091 |

iST Kits compatible with label-free & metabolic labeling; iST-NHS Kits compatible with chemical labeling.

## Ordering information:

http://www.preomics.com/quote

order@preomics.com

## Acknowledgements

We kindly acknowledge early access to the NIST Candidate RM 8461 human liver for proteomics by Benjamin Neely and Clay Davis (National Institute of Standards and Technology, Charleston, SC, USA).

## References

1. Kulak NA, Pichler G, Paron I, Nagaraj N, Mann M. Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells. Nat Methods. 2014 Mar;11(3):319-24. doi: 10.1038/nmeth.2834.

2. Tyanova S, Temu T, Cox J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. Nat Protocols (2016) 11(12):2301-19. doi: 10.1038/nprot.2016.136Application Note: High-throughput homogenization technique for deeper analysis of multiple mouse tissue proteomes. https://www.preomics.com/resources

3. Tyanova S, Temu T, Sinitcyn P, Carlson A, Hein MY, Geiger T, Mann M, Cox J. The Perseus computational platform for comprehensive analysis of (prote)omics data. Nat Methods (2016) 13(9):731-40. doi: 10.1038/nmeth.3901