Article

Selective TiO2 Phosphopeptide Enrichment of Complex Samples in the Nanogram Range

Gábor Tóth 1,2,†, Fanni Bugyi 1,3,†, Simon Sugár 1,4, Goran Mitulović 5,6, Károly Vékey 1,   
Lilla Turiák 1 and László Drahos 1,\*

1 MS Proteomics Research Group, Research Centre for Natural Sciences, 1117 Budapest, Hungary; toth.gabor@ttk.hu (G.T.); bugyi.fanni@ttk.hu (F.B.); sugar.simon@ttk.hu (S.S.); vekey.karoly@ttk.hu (K.V.); turiak.lilla@ttk.hu (L.T.)

2 Department of Inorganic and Analytical Chemistry, Budapest University of Technology and Economics, 1111 Budapest, Hungary

3 Hevesy György PhD School of Chemistry, Eötvös Loránd University, 1117 Budapest, Hungary

4 Ph.D. School of Pharmaceutical Sciences, Semmelweis University, 1085 Budapest, Hungary

5 Proteomics Core Facility, Medical University of Vienna, 1090 Vienna, Austria; goran.mitulovic@meduniwien.ac.at

6 Clinical Department of Laboratory Medicine, Medical University of Vienna, 1090 Vienna, Austria

**\*** Correspondence: drahos.laszlo@ttk.hu; Tel.: +36-1-382-6542

† Shared first authorship.

1. Detailed Investigation of the Selectivity Differences of the Tested Methods

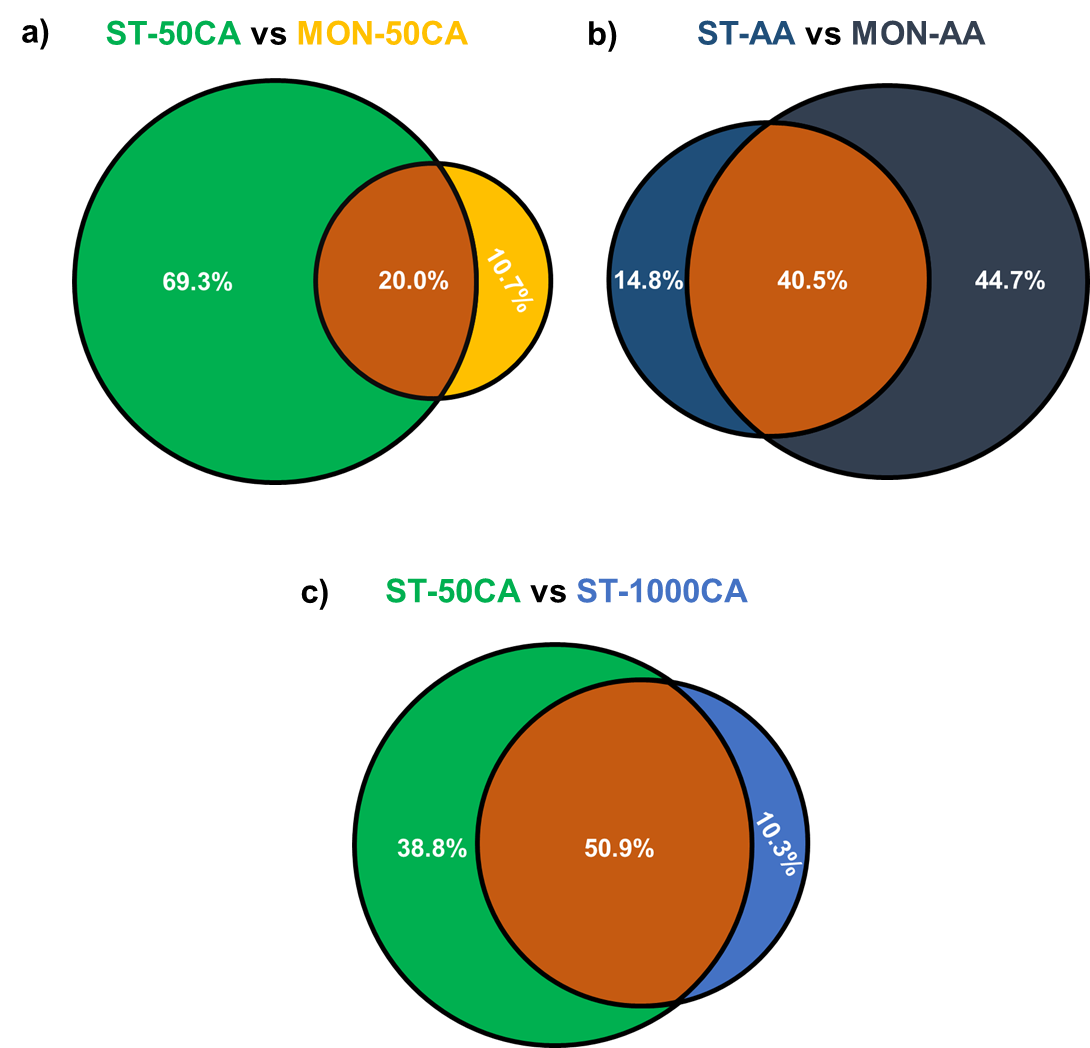
For the corresponding information on this topic, see Section 2.2. in the main body.

When describing the selectivity differences due to the stationary phase, method pairs ST-50CA/MON-50CA (Figure S1a) and ST-AA/MON-AA (Figure S1b) were compared. In the case of the citric acid-based loading buffer, the spin tip performed significantly better (Figure S1a).

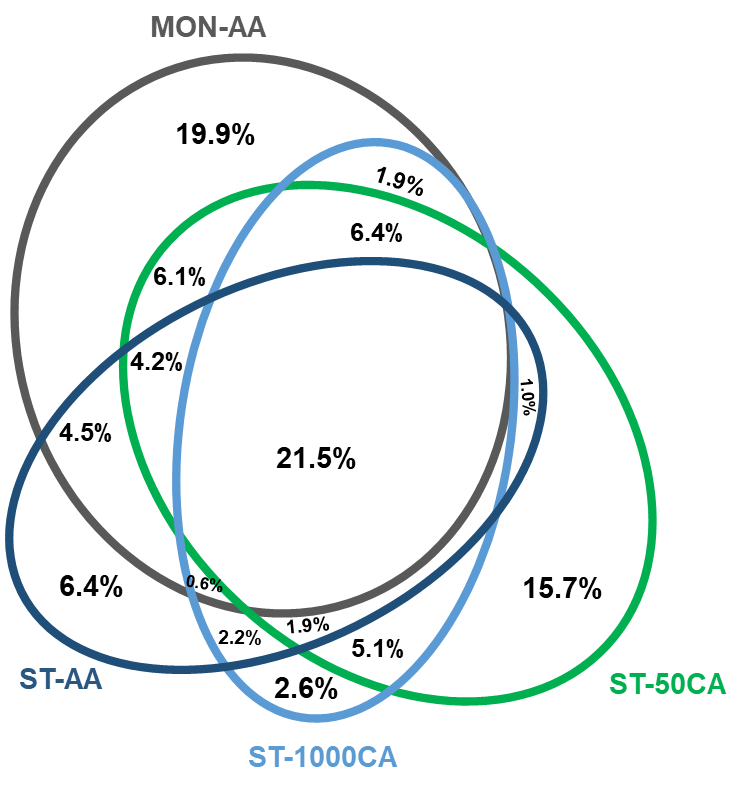
However, when using the acetic acid-based loading buffer, the monolithic column performed better; and a larger intersection, ca. 3/4 of the peptides identified with the spin tip, was observed. Furthermore, a larger portion of PPs identified solely using the monolithic column method was also observed (Figure S1b).

The concentration of the displacing agent influences the column’s selectivity (Figure S1c). While using the ST-50CA and the ST-1000CA methods, the commonly identified PPs only accounted for 50.9% of all the unique PPs, 38.8%, and 10.3% were identified only with the ST-50CA and the ST-1000CA method, respectively. For the entire peptide lists, see Supplementary Table.

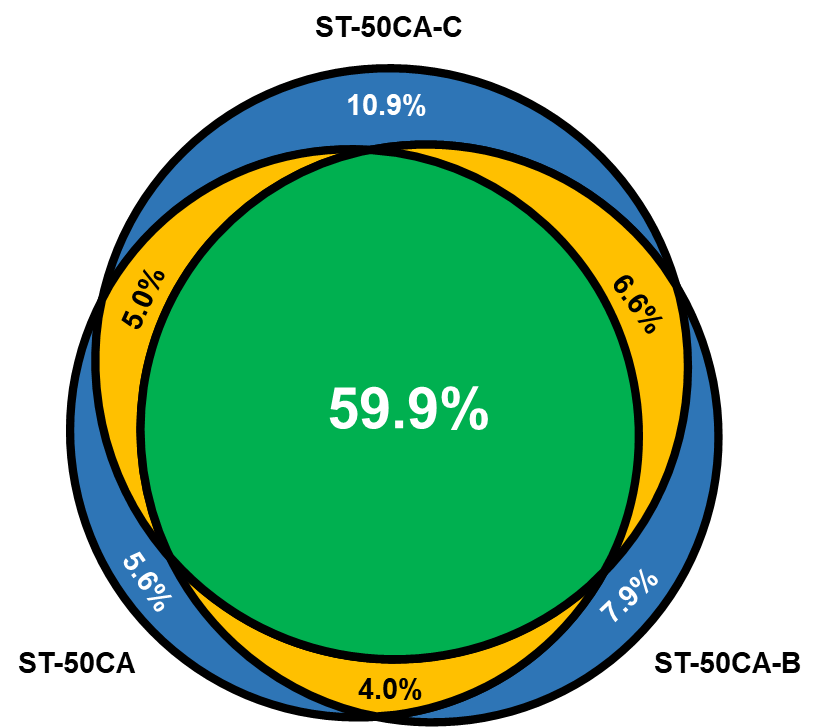
The selectivity differences between the four methods that provided the highest identification rate and conventional peptide length distribution are shown in Figure S2.



**Figure S1.** Selectivity differences between methods comparing the identified PPs. (**a**) Percentage of PPs identified with the ST-50CA and the MON-50CA methods; (**b**) Percentage of PPs identified with the ST-AA and the MON-AA methods; (**c**) Percentage of identified PPs with the ST-50CA and the ST-1000CA methods.



**Figure S2.** Venn diagram comparison of unique PPs identified with the ST-AA, ST-1000CA, ST-50CA, and MON-AA methods.

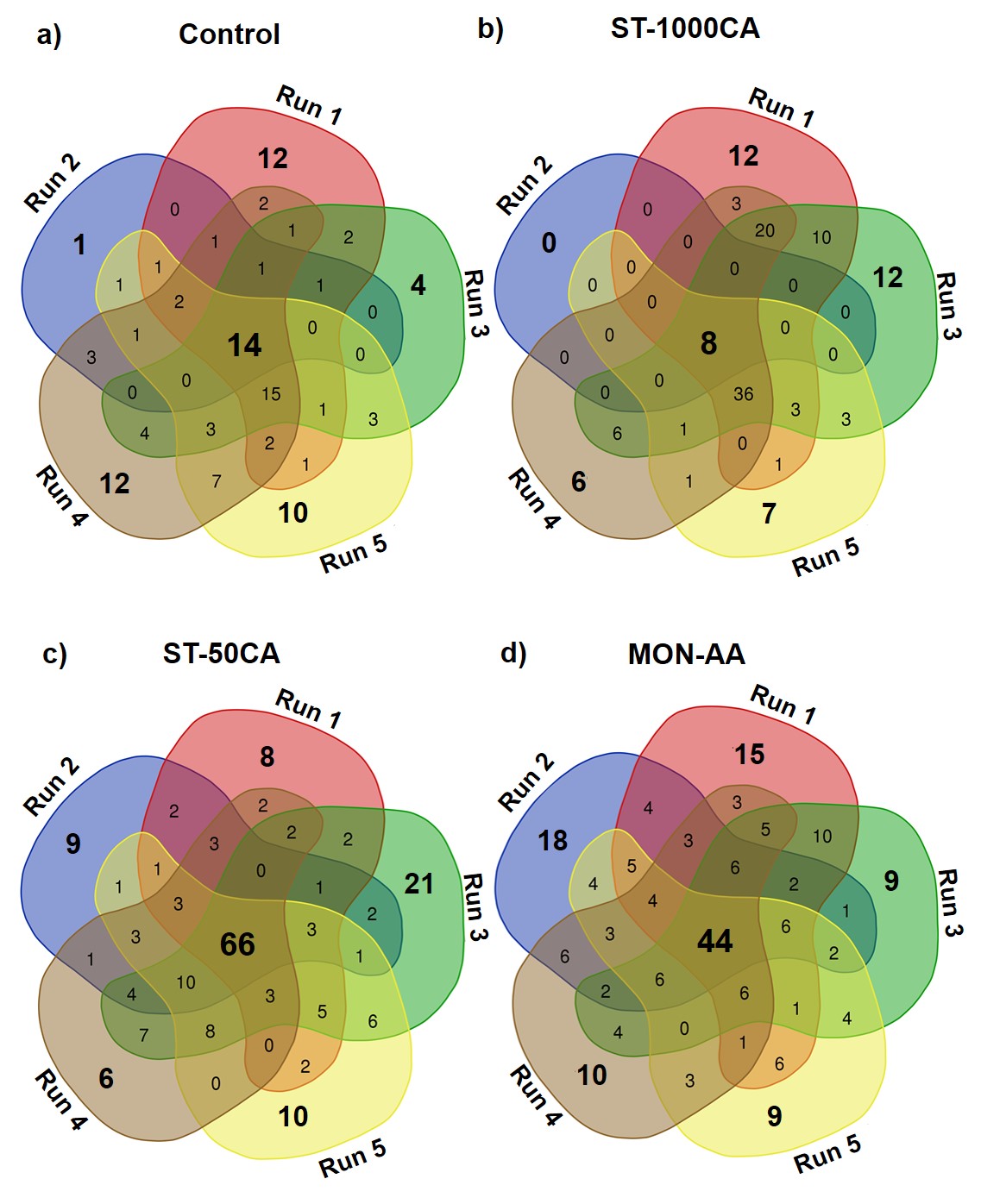


**Figure S3.** Selectivity difference of the modified ST-50CA methods.

2. Repeatability of the Methods

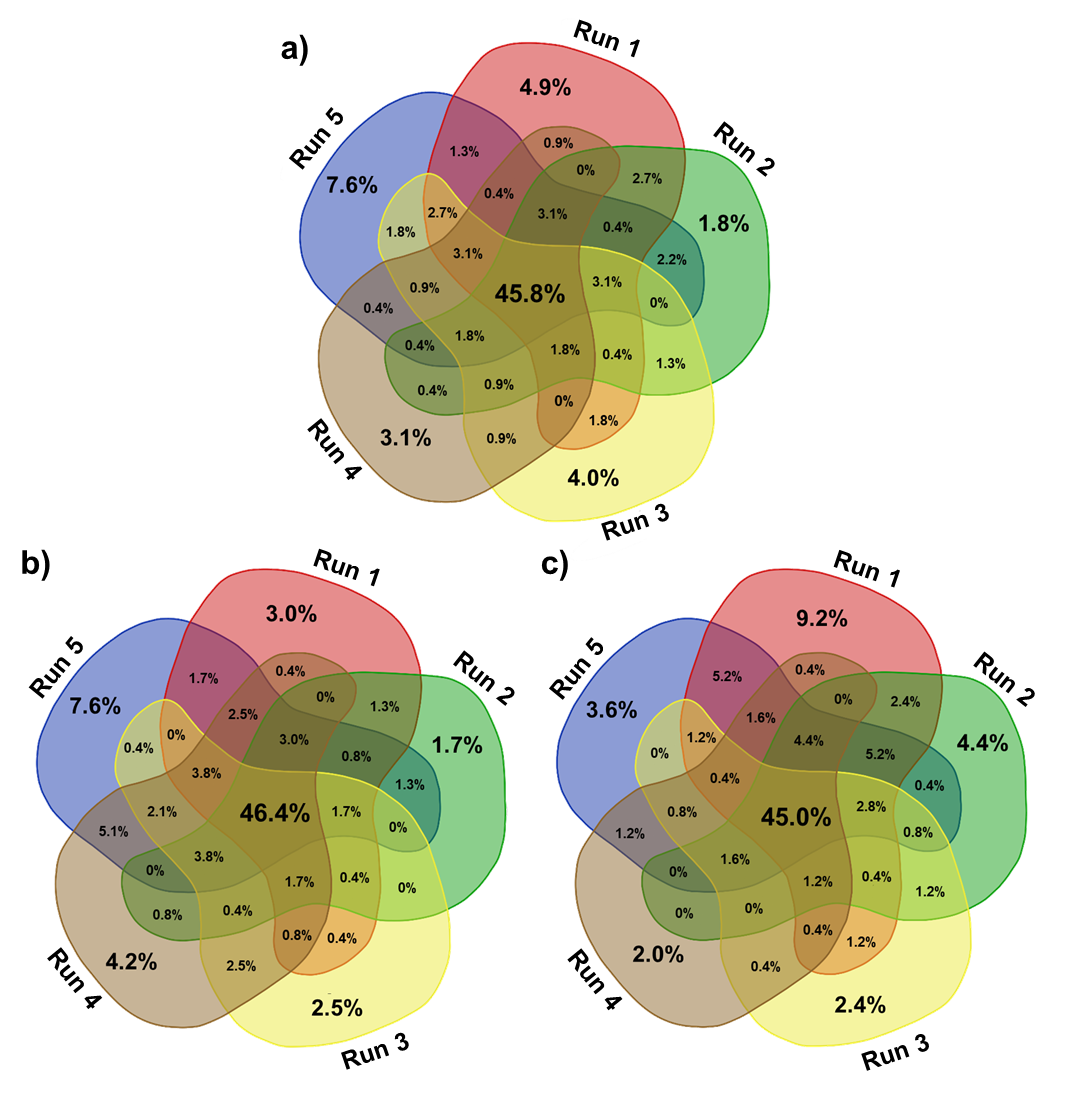
Repeatability of the methods can be assessed from two points of view: the variations in the identification numbers and the quantitative values. In this supplementary section, the qualitative repeatability is assessed in details. For this, we analyzed the PPs identified in five parallel enrichments and five parallels for the control samples. As the control samples were subjected only to the C18 clean-up, the variability seen in Figure S4a is due to the recovery variability of the C18 clean-up (minor) and the identification variability of the data-dependent acquisition (DDA; major). Since for the DDA method, tandem MS is performed on a certain number of the most intensive peaks in a given timeframe, a relatively large variance is observed among parallel runs. In an un-enriched control sample, the majority of the identified compounds are NPs, thus the identification of low-abundance PPs is strongly compromised. In the control sample 13.3% of the unique PPs were identified in all the five runs and 37.1% of them in only one run (note, same distribution can be observed for simple injection parallels).

When proper enrichment is performed, these numbers are closer to each other, since most of the highly abundant NPs are removed resulting in better identification repeatability of the PPs. ST-1000CA poses an example for irreproducible enrichment (Figure S4b): 6.1% of the PPs were identified in all five runs, while most of the PPs were identified in 3 or 4 runs, and 28.2% of the PPs in only 1 run. This means that although a large portion of the NPs was removed during the enrichment (see Table 2 in Paragraph 3.1. of the main body), the binding of PPs to the MOAC stationary phase is not satisfactory and not repeatable. The two best-performing methods regarding identification performance had outstanding repeatability as well. With the ST-50CA method, 2.6-times larger ratio of the PPs was identified in all five runs as compared to the control, while this increase was 1.6-times for the MON-AA method (Figure S4c,d). In parallel, the ratio of PPs identified in only one run decreased to 28.1% (ST-50CA) and 30.2% (MON-AA). This implies that decreasing sample complexity significantly reduced the variance of the DDA and additional sample preparation steps did not increase it.



**Figure S4.** Qualitative (identification) repeatability of (**a**) the Control (variance of DDA method), (**b**) the ST-1000CA method, (**c**) the ST-50CA method, and (**d**) the MON-AA method. For each method, 5 parallel enrichments were performed.

The qualitative repeatability of the improved ST-50CA methods was also addressed in a similar way. For the discussion, please see Section 2.5. in the main body.



**Figure S5.** Qualitative repeatability of the modified 50 mM citric acid-based methods. (**a**) ST-50CA, (**b**) ST-50CA-B, (**c**) ST-50CA-C.