Supporting Information

S.1 Retention times and peak widths of test proteins analyzed with RPLC and HILIC

S.2 Scouting gradient analyses of *IdeS*-digested trastuzumab and intact glycoproteins

S.3 Parameters of tested glycoproteins for different retention models

S.4 Optimization of the separation protein glycoforms using PIOTR

S.5 Gradient retention factors of the scouting gradients and optimized methods for the tested glycoproteins

S.6 Extracted-ion chromatograms of the peaks of the Fc portion of trastuzumab and Ovalbumin analyzed by HILIC-MS

S.1 Retention times and peak widths of test proteins analyzed with RPLC and HILIC

Table S1. Retention time (RT), peak width at half height, experimentally-obtained coefficient S and Akaike Information Criterion (AIC) of the test proteins obtained after RPLC and HILIC analysis.

			RPLC				**HILIC**			
Protein	MW (Da) [1]	UniProt entry nr.	RT (min)	Peak width at half height (min)	Coefficient S	AIC	RT (min)	Peak width at half height (min)	Coefficient S	AIC
RnB	13690	P61823	8.53	0.19	22.9	-30.6	21.07	0.31	22.8	-19.6
RnA	13690	P61823	8.73	0.21	21.1	-33.5	17.10	0.30	25.8	-13.0
cyt C	11702	P00004	12.68	0.24	26.8	-29.2	12.56	0.32	31.1	-22.7
ubi	8565	P62992	13.17	0.23	25.3	-35.2	8.49	0.24	36.9	-22.4
lys	14313	P00698	14.61	0.17	30.1	-33.1	12.90	0.24	29.7	-34.8
tryp	23305	P00760	15.86	0.24	34.2	-40.6	16.22	0.33	32.5	-13.4
trans	75091	Q06AH7	16.38	0.27	48.2	-33.9	19.98	0.34	32.9	-25.8
AGP	21560	P02763	16.59	0.55	46.0	-28.8	23.53	2.40	20.4	-18.6
BSA	66433	P02769	17.57	0.31	66.4	-26.6	15.79	0.30	41.1	-23.2
myo	16951	P68082	19.29	0.21	52.4	-27.1	10.77	0.23	42.2	-29.6
c.tryp	25244	P00766	19.54	0.32	42.4	-34.9	12.84	0.37	32.5	-20.1
fet	40846	Q58D62	20.58	0.55	47.4	-44.0	25.19	1.03	25.4	-25.4
CA	28983	P00921	20.87	0.20	57.4	-31.3	14.61	0.23	36.0	-13.2
ova	42750	P01012	24.15	0.51	96.3	-28.5	12.59	1.07	39.4	-18.7
thyro	301219	P01267	24.74	1.32	160.3	-18.4	23.38	1.52	16.8	-18.5

[1] Molecular mass (MW) was calculated from the amino acid sequence without considering proteoforms or dimerization.

Table S2. RPLC elution order and the number of hydrophobic amino acid residues of the non-glycosylated test proteins. Between brackets the percentage with respect to the total number of amino acids of the protein.

Protein	Elution order	Ala	Ile	Leu	Met	Phe	Val	Pro	Gly	Sum	Total residues
RnA	1	12 (9.7)	3 (2.4)	2 (1.6)	4 (3.2)	3 (2.4)	9 (7.3)	4 (3.2)	3 (2.4)	28 (22.6)	124
cyt C	2	6 (5.8)	6 (5.8)	6 (5.8)	2 (1.9)	4 (3.8)	3 (2.9)	4 (3.8)	12 (11.5)	43 (41.3)	104
ubi	3	2 (2.6)	7 (9.2)	9 (11.8)	1 (1.3)	2 (2.6)	4 (5.3)	3 (3.9)	6 (7.9)	34 (44.7)	76
lys	4	12 (9.3)	6 (4.7)	8 (6.2)	2 (1.6)	3 (2.3)	6 (4.7)	2 (1.6)	12 (9.3)	51 (29.5)	129
tryp	5	14 (6.3)	15 (6.7)	14 (6.3)	2 (0.9)	3 (1.3)	17 (7.6)	8 (3.6)	25 (11.2)	98 (43.9)	223
BSA	6	47 (8.1)	14 (2.4)	61 (10.5)	4 (0.7)	27 (4.6)	36 (6.2)	28 (4.8)	16 (2.7)	233 (40.0)	583
myo	7	15 (9.8)	9 (5.9)	17 (11.1)	2 (1.3)	7 (4.6)	7 (4.6)	4 (2.6)	15 (9.8)	76 (49.7)	153
c.tryp	8	22 (9.0)	10 (4.1)	19 (7.8)	2 (0.8)	6 (2.5)	23 (9.4)	9 (3.7)	23 (9.4)	114 (46.7)	244
CA	9	17 (6.6)	5 (1.9)	26 (10.0)	3 (1.2)	11 (4.2)	20 (7.7)	19 (7.3)	20 (7.7)	121 (46.7)	259

Table S3. HILIC elution order and the number of hydrophilic amino acid residues of the non-glycosylated test proteins. Between brackets the percentage with respect to the total number of amino acids of the protein.

Protein	Elution order	Gln	Asp	His	Ser	Thr	Tyr	Cys	Trp	Sum	Total residues
ubi	1	6 (7.9)	2 (2.6)	1 (1.3)	3 (3.9)	7 (9.2)	1 (1.3)	0 (0.0)	0 (0.0)	20 (26.3)	76
myo	2	6 (3.9)	2 (1.3)	11 (7.2)	5 (3.3)	7 (4.6)	2 (1.3)	0 (0.0)	2 (1.3)	35 (22.9)	153
cyt C	3	3 (2.9)	5 (4.8)	3(2.9)	0 (0.0)	10 (9.6)	4 (3.8)	2(1.9)	1 (1.0)	28 (26.9)	104
c.tryp	4	10 (4.1)	14 (5.7)	2 (0.8)	28 (11.5)	23 (9.4)	4 (1.6)	10 (4.1)	8 (3.3)	99 (40.6)	244
lys	5	3 (2.3)	14 (10.9)	1 (0.8)	10 (7.8)	7 (5.4)	3 (2.3)	8 (6.2)	6 (4.7)	52 (40.3)	129
CA	6	12 (4.6)	13 (5.0)	11 (4.2)	16 (6.2)	14 (5.4)	8 (3.1)	0 (0.0)	7 (2.7)	81 (31.3)	259
tryp	7	10 (4.5)	16 (7.2)	3 (1.3)	34 (15.2)	10 (4.5)	10 (4.5)	12 (5.4)	4 (1.8)	99 (44.4)	223
BSA	8	20 (3.4)	14 (2.4)	17 (2.9)	28 (4.8)	33 (5.7)	20 (3.4)	35 (6.0)	2 (0.3)	169 (29.0)	583
RnA	9	7 (5.6)	10 (8.1)	4 (3.2)	15 (12.1)	10 (8.1)	6 (4.8)	8 (6.5)	0 (0.0)	60 (48.4)	124

S.2 Scouting gradient analyses of *IdeS*-digested trastuzumab and intact glycoproteins

Scouting gradients of *IdeS*-digested trastuzumab (Figure S1), fet (Figure S2), AGP (Figure S3), and ova (Figure S4).

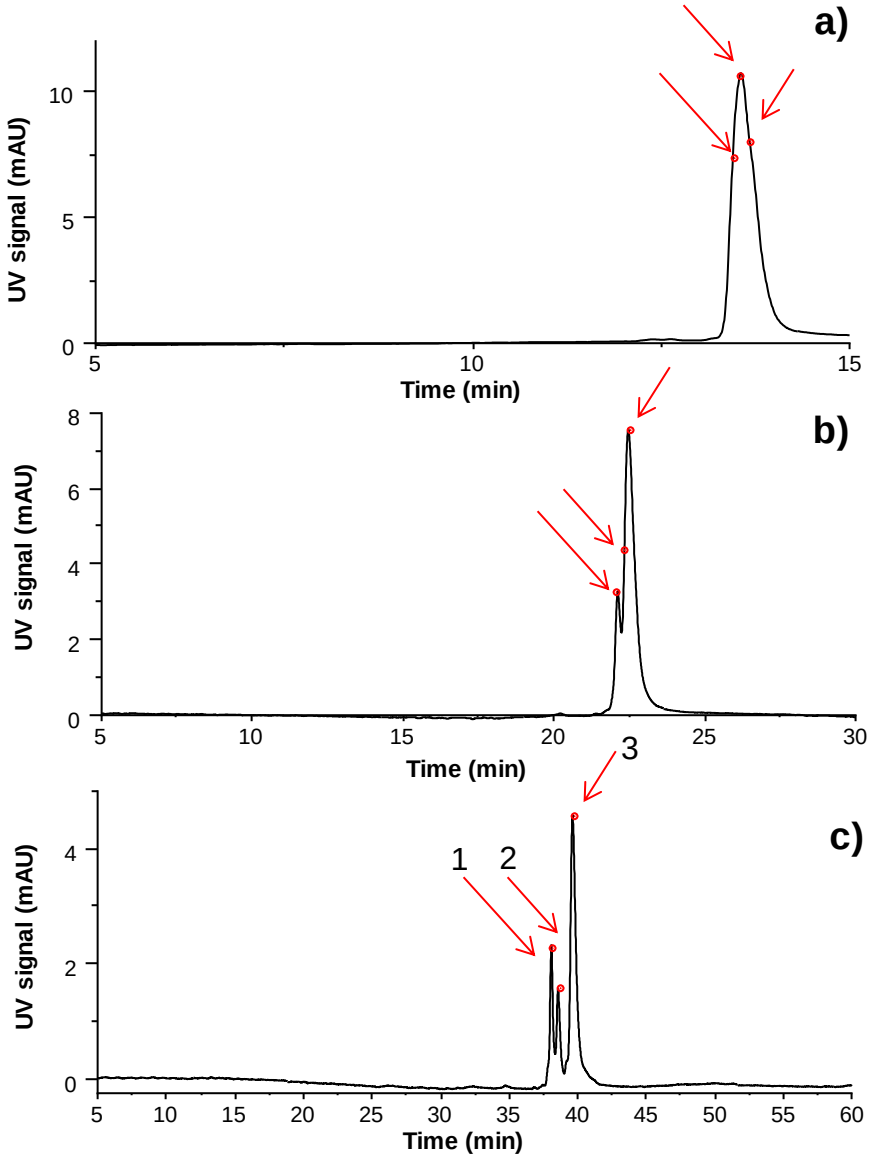

Figure S1. HILIC-UV of *IdeS*-digested trastuzumab using a linear gradient from 10 to 50% B in (a) 15, (b) 30 and (c) 60 min. The red dots and arrows indicate the features (retention times) used for modeling and optimization by PIOTR. Three features (peak maxima 1-3) were chosen at t_G = 60 min). At t_G = 30 min, two features (peak maxima) were selected; the third feature was chosen at half of the distance between the peak maxima. At t_G = 15 min no clear features showed and the peak maximum and the points corresponding to the peak width at half height (front and tail) of the peak were assigned as features.

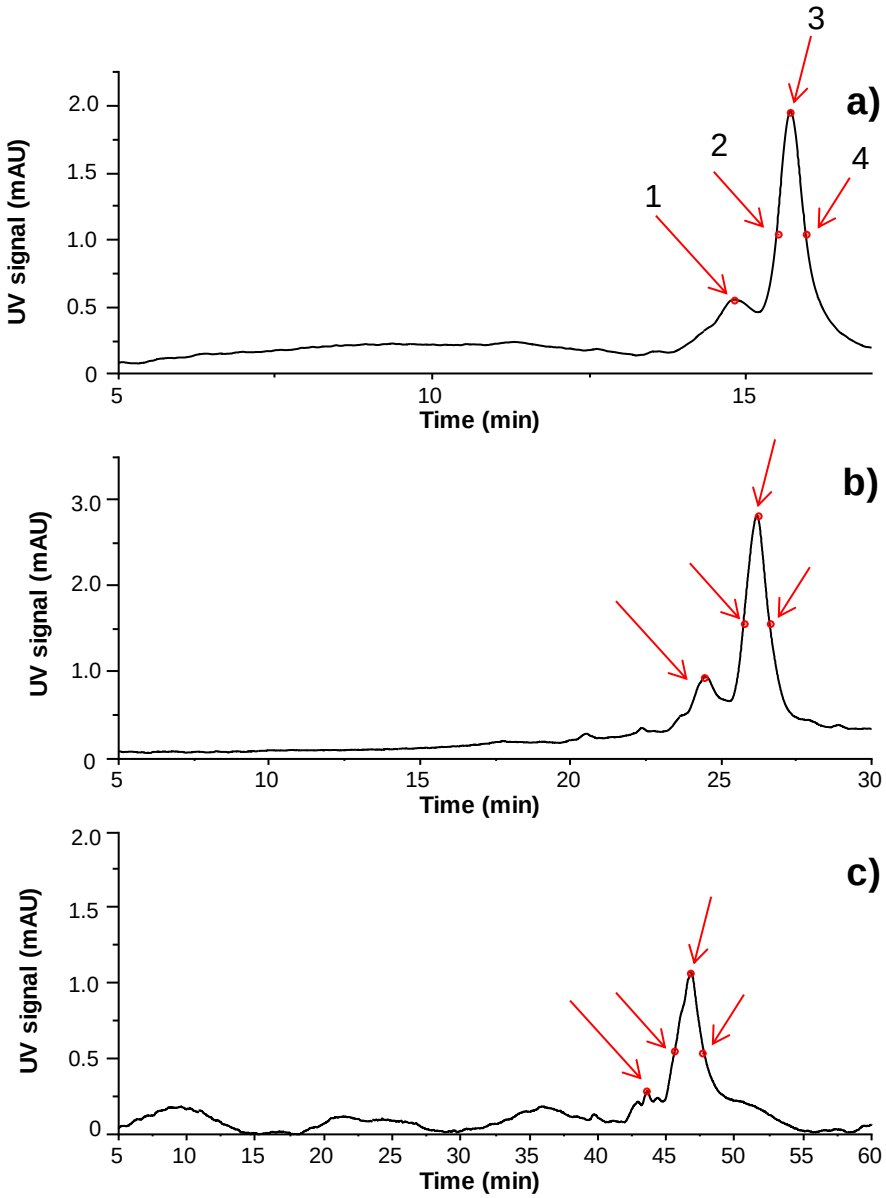

Figure S2. HILIC-UV of fet using a linear gradient from 10 to 50% B in (a) 15, (b) 30 (b) and (c) 60 min. The red dots and arrows indicate the features (retention times) used for modeling and optimization by PIOTR. Four features were selected. One corresponded to the maximum of the first eluting peak (1) and the other to the peak maximum of the second peak (3) and the point corresponding to the peak width at half height (front (2) and tail (4) of the second peak).

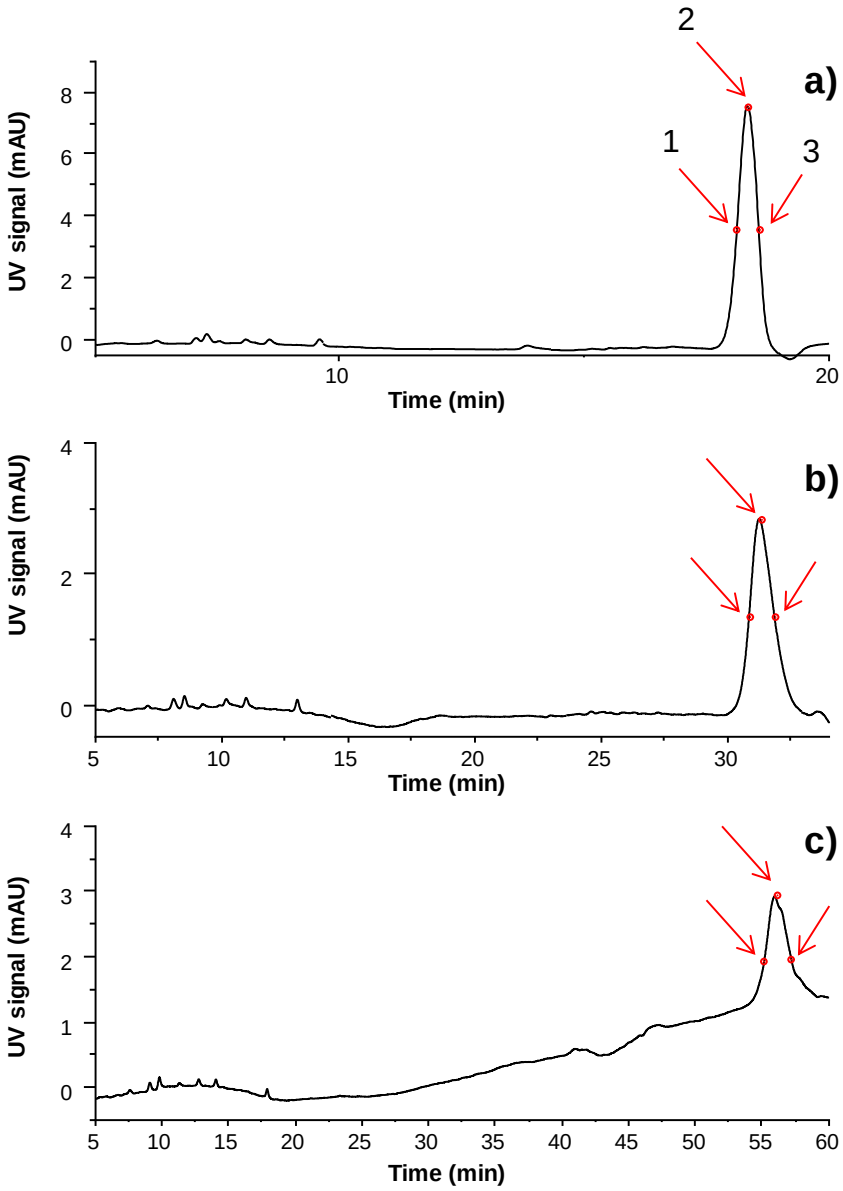

Figure S3. HILIC-UV of AGP using a linear gradient from 10 to 50% B in (a) 15, (b) 30 (b) and (c) 60 min. The red dots and arrows indicate the features (retention times) used for modeling and optimization by PIOTR. Three features were selected corresponding to the peak maximum (2) and to the peak width at half height (front (1) and tail (3).

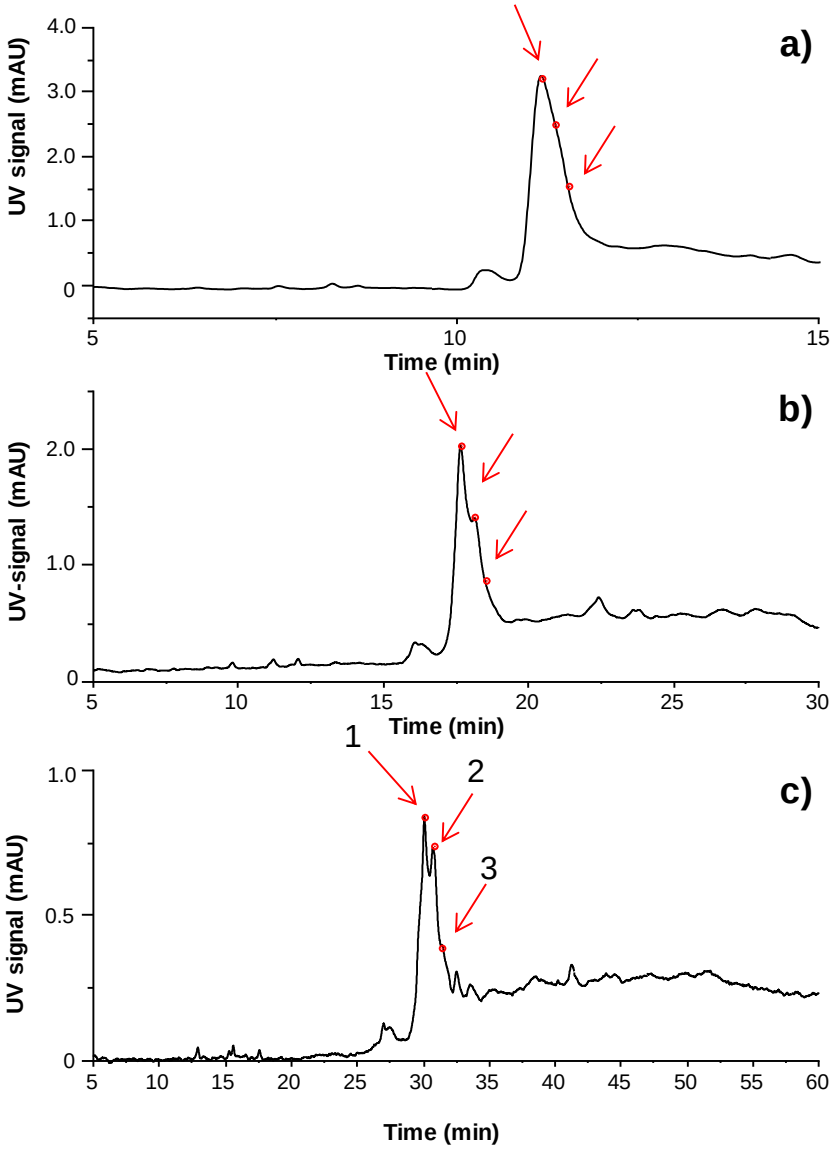

Figure S4. HILIC-UV of ova using a linear gradient from 10 to 50% B in (a) 15, (b) 30 and (c) 60 min. The red dots and arrows indicate the features (retention times) used for modeling and optimization by PIOTR. Three features (peak maxima 1-3) were chosen at t_G = 60 min. These features were then tentatively assigned in the shorter gradient times (a and b).

S.3 Parameters of tested glycoproteins for different retention models

Table S4 lists the parameters obtained for the following retention models: mixed-mode, Neue-Kuss, quadratic, adsorption and LSS. The AICs classified in five ranges are reported in Figure 4 of the manuscript.

The equation and parameters of the LSS model is given in the manuscript (Equation 4). The equations and parameters of the other models are given below. $ln\ k$ is the natural logarithm of the retention factor, φ is the volume fraction of the strong solvent in a binary eluent, and k_0 is the extrapolated (not necessarily real) k of the analyte in pure weak solvent (i.e. φ equals 0) [24].

The adsorption model describes the relationship between the natural logarithm of the retention factor (k) and the natural logarithm of the composition φ (Equation S1), where n is the ratio of surface areas occupied by water and by solute molecules [21].

$$ln\ k = ln\ k_0 - n ln\ \varphi \qquad \text{Equation S1}$$

The quadratic model [42] is a nonlinear model, where S_1 is the slope, and S_2 is the curvature coefficient (Equation S2).

$$ln\ k = ln\ k_0 + S_1\varphi + S_2\varphi \qquad \text{Equation S2}$$

The Neue-Kuss model gives a mathematical description of the curved relationship between the natural logarithm of the retention factor and the solvent composition, where S_1 is the slope, and S_2 is the curvature coefficient (Equation S3) [26].

$$ln\ k = ln\ k_0 + 2\ ln(1 + S_1\varphi) - \frac{S_1\varphi}{1+S_2\varphi} \qquad \text{Equation S3}$$

The mixed-mode model was proposed to describe the retention in HILIC [25], where S_1 accounts for the interaction of analytes with the stationary phase and S_2 for the interaction of analytes with the solvents (Equation S4).

$$ln\ k = ln\ k_0 + S_1\varphi + S_2 ln\ \varphi \qquad \text{Equation S4}$$

Table S4. The parameters of the different retention models and the AIC values obtained.

Model		*Mixed-mode*				*Neue-Kuss*				*Quadratic*				*Adsorption*			*LSS*		
Protein standard	**Peak**	**ln k0**	**S1**	**S2**	**AIC**	**ln k0**	**S1**	**S2**	**AIC**	**ln k0**	**S1**	**S2**	**AIC**	**ln k0**	**n**	**AIC**	**ln k0**	**S**	**AIC**
Ovalbumin	1	5.2	30.6	3.3	-17.1	9.8	20.9	-1.1	-14.9	13.3	-46.9	8.1	-16.8	-15.3	12.3	-20.1	13.3	46.2	-18.6
	2	2.3	26.6	4.8	-2.6	10.0	20.9	-1.1	-2.6	13.4	-45.4	3.3	-2.6	-15.9	12.9	-4.6	13.9	47.5	-4.6
	3	4.6	26.8	3.3	10.4	9.6	20.4	-1.0	10.2	12.1	-38.4	0.4	10.3	-13.8	11.6	8.6	12.7	41.6	8.3
Fetuin	1	0.8	17.1	6.7	-11.3	7.7	5.9	-1.6	-5.5	14.3	-35.9	1.1	-10.4	-12.9	13.5	-14.1	14.7	38.2	-12.3
	2	0.3	15.9	7.4	-15.7	5.7	-0.1	-2.4	-3.4	28.7	-106.3	90.7	-20.2	-12.8	14.0	-18.8	15.2	38.0	-16.1
	3	1.7	17.5	6.8	-12.3	8.0	5.9	-1.5	-6.5	16.1	-41.5	8.6	-11.8	-12.7	14.1	-15.1	15.3	37.7	-13.4
	4	2.3	18.0	6.6	-12.7	6.4	1.5	-1.9	-1.0	17.1	-45.7	14.2	-12.3	-12.5	14.2	-15.8	15.4	37.3	-13.5
AGP	1	2.8	14.6	6.5	-17.5	2.0	7.5	5.1	83.0	16.3	-37.1	9.0	-16.8	-10.0	13.5	-21.1	14.9	31.3	-17.6
	2	2.4	14.0	6.8	-15.9	6.7	1.8	-1.6	-1.9	17.8	-42.9	15.3	-15.8	-9.9	13.6	-19.3	15.0	31.2	-16.2
	3	1.8	13.5	7.6	-8.1	5.0	-0.4	-2.0	31.1	15.4	-31.3	1.9	-7.8	-10.0	13.8	-11.2	15.2	31.4	-9.9
Thyroglobulin	1	0.7	30.3	13.0	-5.3	9.5	5.1	-1.8	-2.5	31.7	-92.6	36.8	-5.3	-23.9	25.3	-7.6	17.7	42.5	0.5
	2	-0.8	24.5	12.5	-11.1	9.8	6.3	-1.7	-7.3	24.2	-60.8	6.1	-10.7	-21.0	22.7	-13.5	16.7	39.3	-3.0
	3	-4.8	17.4	14.2	-10.7	2.1	6.9	4.5	78.3	22.5	-53.8	1.8	-10.2	-19.6	21.6	-13.0	16.3	37.7	-3.3
IdeS-digested Trastuzumab	1	0.6	19.5	5.3	2.6	12.4	28,1	-0.5	2.8	25.6	-105.8	97.9	2.5	-14.2	13.5	0.6	14.6	41.7	0.7
	2	1.0	22.3	6.3	4.2	13.4	29.8	-0.6	4.5	20.2	-69.3	38.2	4.2	-16.0	15.2	2.1	16.2	46.4	2.3
	3	7.1	37.6	6.1	8.9	15.9	32.3	-0.7	9.2	47.4	-210.2	220.8	8.3	-21.2	20.1	6.7	21.0	60.0	7.0

S.4 Optimization of the separation of protein glycoforms using PIOTR

First, three scouting gradients were measured (Figure 3 in manuscript and Figures S1-4). The scouting gradients were used to generate the retention time of the main peak and (partially) resolved glycoforms. For each gradient condition, we tabulated three retention times (4 in the case of fetuin) and these were imported in PIOTR as retention times belonging to distinct analytes. Below we briefly describe the procedure to select the retention times at different gradient times.

- ***IdeS*-digested trastuzumab**: The retention time for *IdeS*-digested trastuzumab were derived from the three distinct features observed with a gradient time of 60 min (indicated as feature 1-3 in the figure S1). In the 30-min gradient, we observed only 2 features (1 and 3), the third retention time was assigned at half of the distance between the two. In the 15-min gradient, no clear feature was observed and the peak maximum was used as retention time of the second feature (2) and the retention times for feature 1 and 3 were assigned based on their relative distance respect to feature 2, previously observed in longer scouting gradients.
- **Ova**: For this protein, we collected the retention times of three of the features observed with the gradient with the longest gradient time. These were then tentatively assigned in the shorter gradient times.
- **Fet, AGP, and thyro**: These proteins did not present resolved species also when using long gradient times (except for Fet that presented two main peaks in all the analysis). From this analysis, we extracted three retention times, corresponding to the peak maximum (indicated as 2) and to the peak width at half height (beginning (1) and end (2) of the peak).

The retention times for each of the features under different gradient times were imported in PIOTR (together with dwell and dead volumes). Mixed-mode, Neue-Kuss, Quadratic, Adsorption and LSS models were used to fit to fit the data obtained for each feature. The MATLAB-based PIOTR program (Program for Interpretive Optimization of Two-dimensional Resolution) used the FMINSEARCH function of MATLAB to solve the integrated gradient equations to determine the retention parameters for each compound. The resulting retention parameters are shown in Table S-4.

Table S-4 also lists the Akaike Information Criterion (AIC) for each compound. This goodness-of-fit value indicates how well the used retention model can describe the retention behavior. Generally, lower values reflect a better ability of the retention model to describe the retention. The analysis of the AIC value is discussed in Figure 4 of the manuscript. As discussed above, the adsorption process was used for the optimization process.

After that, possible HILIC methods were created based on calculated retention parameters by varying three different factors: starting percentage solvent B (φ_{init}), gradient time (up to 60 min), and final percentage solvent B (φ_{final}). First, broad ranges for solvent B with a low number of increments (steps) of %B were used to get an indication of the optimal conditions. In this case, the φ_{init} was from 0.20 to 0.45 in 10 steps of 2.5%, the φ_{final} was from 0.25 to 0.50 in 10 steps of 2.5%, and the gradient time was from 15 to 60 min in 45 steps of 1 min. The number of possible options in this case is 10 (i.e. the number of steps). Then the optimization factors were further specified per protein depending on the predicted methods, as is reflected in Table S5. The number of potential methods is the product of all options. For instance, for thyro the total number of predicted methods was 4320.

Table S5. Specified ranges and steps of each glycoprotein, where φ_{init} is the starting percentage of solvent B, φ_{final} is the final percentage of solvent B, and t_G is the gradient time in min.

	$\varphi_{\mathbf{init}}$			$\varphi_{\mathbf{final}}$			t_G **(min)**		
	Min. value	Max. value	Num. Steps	Min. value	Max. value	Num. Steps	Min. value	Max. value	Num. Steps
trastuzumab	0.23	0.33	20	0.25	0.35	20	20	60	40
ova	0.20	0.30	20	0.25	0.30	20	30	60	30
fet	0.25	0.35	20	0.30	0.35	20	15	60	45
AGP	0.35	0.40	20	0.40	0.50	20	15	60	45
thyro	0.32	0.35	12	0.39	0.42	12	30	60	30

In Figure S5, all the simulated methods for thyro are summarized. The retention times were predicted for each of the generated methods and the quality descriptor resolution was calculated according to Schure's metric for resolution [43]. The software generated peak width values at different gradient times using Van Deemter parameters (A: 2; B: 3; C: 0.25). The resolutions between all the peaks were summarized into a Deringer-based score [44]. In Figure S5, we report the resolution values plotted against the analysis time (time for the last analyte to elute) and the gradient time in Pareto plots. Of all the solution reported we selected the one having highest value of resolution (i.e. resolution score is 1), in which the peaks eluted within the gradient time, and having the lowest gradient time in the interval between 40 and 55 min (with a maximum total analysis time of 70 min). The chosen method is indicated in red. Table S6 provides the predicted retention time and %B at the moment of elution per model and the obtained experimental retention time of *IdeS*-digested trastuzumab. Table S7 reports the

difference between calculated and experimental retention time and %B of the adsorption and LSS model.

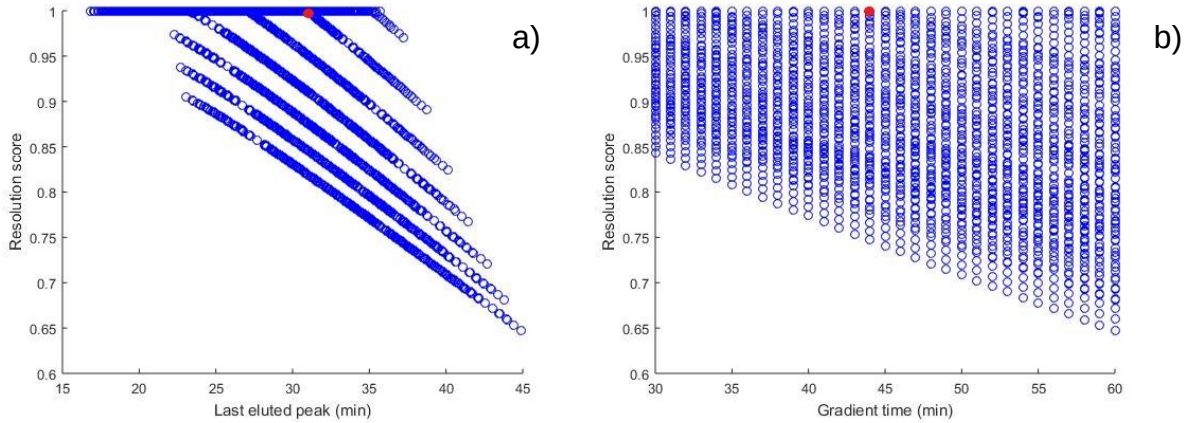

Figure S5. Pareto plot of (a) resolution score vs. last eluted peak and (b) resolution score vs. gradient time of thyro. The specified ranges used to obtain these figures are mentioned in Table S5. The chosen optimized method is indicated with a red dot.

Table S6. Comparison of the experimentally obtained retention times of the fragments of trastuzumab and predicted retention times using the different retention models with the requirement of resolution1 and gradient time of 53 min.

Model	**Retention times of the peaks (min)**			**%B of elution of the peaks (%)**		
	Fc/2		**F(ab)'2**	**Fc/2**		**F(ab)'2**
Adsorption	27.8	29.6	37.3	35.4	35.7	36.8
LSS	27.3	29	36.5	35.4	35.6	36.7
mixed mode	32.6	34.8	44.5	36.1	36.4	37.8
Quadratic	32.3	34.3	44.8	36.1	36.4	37.8
Neue-Kuss	27.1	28.7	36.1	35.3	35.6	36.6
Experimental	28.0	30.3	44.7	35.5	35.8	37.8

Table S7. Difference between calculated retention time and %B of elution and the experimentally obtained data reported in Table S6.

Model	**Retention times of the peaks (min)**			**%B of elution of the peaks (%)**		
Adsorption	-0.2	-0.7	-7.4	-0.1	-0.1	-1.0
LSS	-0.7	-1.3	-8.2	-0.1	-0.2	-1.1
HILIC	4.6	4.5	-0.2	0.6	0.6	0
Quadratic	4.3	4	0.1	0.6	0.6	0
Neue-Kuss	-0.9	-1.6	-8.6	-0.2	-0.2	-1.2

S.5 Gradient retention factors of the scouting gradients and optimized methods for the tested glycoproteins

Table S8 depicts the k^* values for the scouting gradients and the optimized methods for the glycoproteins (ova, fet, thryo, AGP, and *IdeS*-digested Trastuzumab).

Table S8. The gradient retention factors (k^*) of the scouting gradients and the optimized method for every chosen retention time of the glycoproteins.

Glycoprotein	**Peak**	**Coefficient S**	k^{*1} **(t_G= 15 min)**	k^{*1} **(t_G= 30 min)**	k^{*1} **(t_G= 60 min)**	k^{*1} **Optimized**
ova	1	46.2	0.45	0.90	1.81	7.59
	2	47.5	0.44	0.88	1.76	7.38
	3	41.6	0.50	1.00	2.01	8.43
fet	1	38.2	0.55	1.09	2.19	8.65
	2	38.0	0.55	1.10	2.20	8.69
	3	37.7	0.55	1.11	2.22	8.76
	4	37.3	0.56	1.12	2.24	8.86
thyro	1	42.5	0.49	0.98	1.97	8.30
	2	39.3	0.53	1.06	2.13	8.97
	3	37.7	0.55	1.11	2.22	9.35
AGP	1	31.3	0.67	1.34	2.67	15.64
	2	31.2	0.67	1.34	2.68	15.69
	3	31.4	0.67	1.33	2.66	15.59
IdeS-digested Trastuzumab	1	41.7	0.50	1.00	2.01	9.51
	2	46.4	0.45	0.90	1.80	8.54
	3	60.0	0.35	0.70	1.39	6.61

[1] Calculated using Equation 3.

S.6 Extracted-ion chromatograms of the peaks of the Fc portion of trastuzumab and Ovalbumin analyzed by HILIC-MS

The extracted-ion chromatograms (Figure S6) of the fragments of *IdeS*-digested trastuzumab and ovalbumin (Figure S7) analyzed by HILIC-MS are provided.

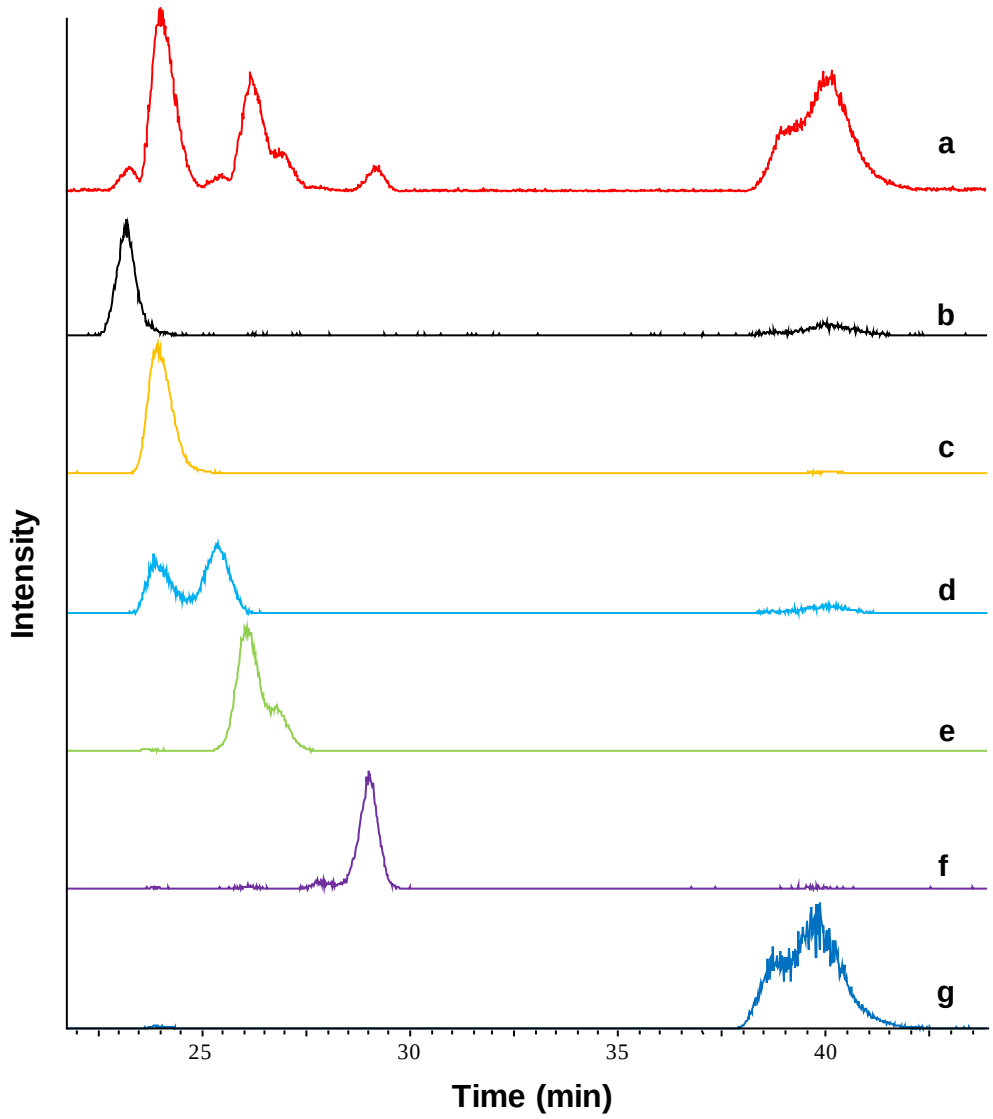

Figure S6. Base peak chromatogram (a) and extracted-ion chromatograms (EICs) of the HILIC-MS separation of the Fc portion of trastuzumab reported in Figure 6 of the manuscript. The extracted ion currents were obtained summing intensities from 3 charge states: (b) m/z 1792.9, 2509.5 and 3136.7, (c) m/z 2804.6, 2524.1 and 3154.9, (d) m/z 2525.7, 2806.1 and 3156.9, (e) m/z 2822.5, 3175.2 and 2540.3, (f) m/z 2840.6, 3195.4 and 2556.6, and (g) m/z 2570.2, 2639.6 and 3616.9. The assignment of the glycoform follows what is described in Figure 6.

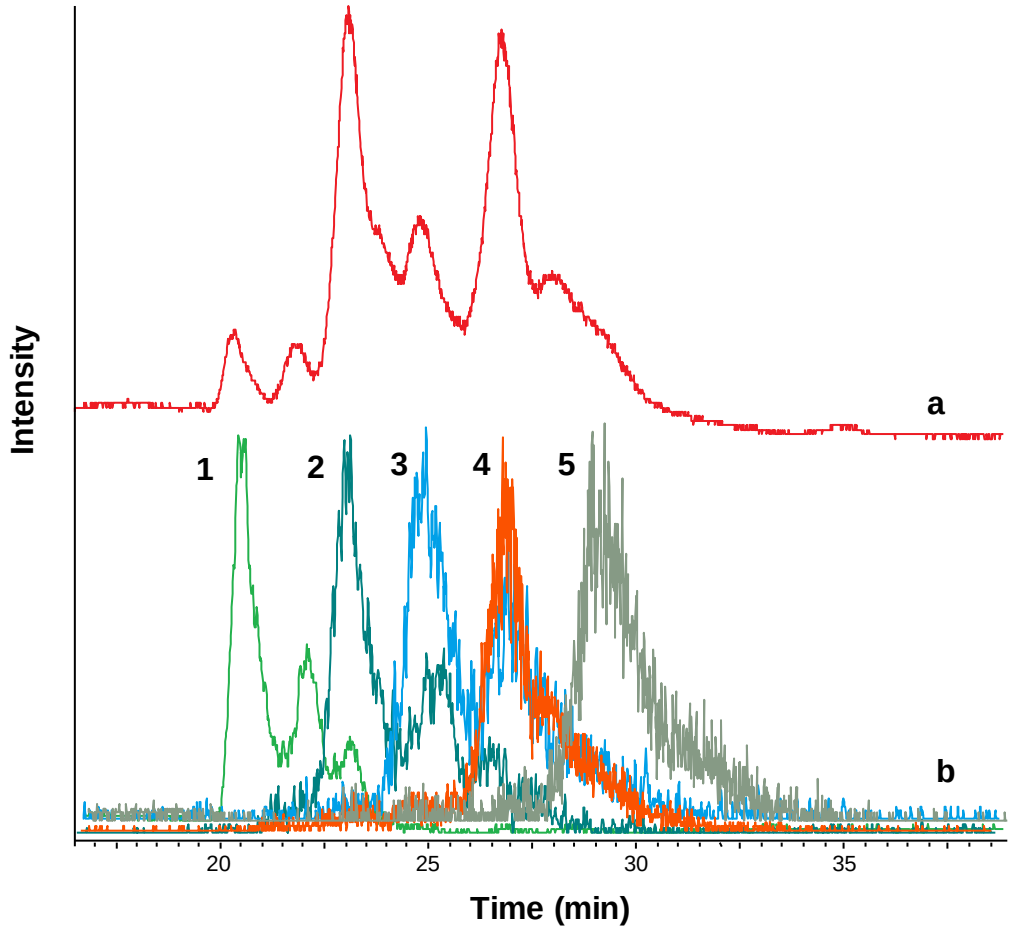

Figure S7. Total Ion Current (a) and extracted-ion chromatograms (EICs) of the HILIC-MS separation of Ovalbumin. The extracted ion currents were done monitoring: (1) m/z 1003.0, 1026.3, 1050.6 , (2) m/z 1030.0, (3) m/z 1065.2, (4) m/z 1088.2, (5) m/z 1093.1