Journal of Chromatography B, 915–916 (2013) 57–63

Contents lists available at SciVerse ScienceDirect

Journal of Chromatography B

journal homepage: www.elsevier.com/locate/chromb

The effect of various S-alkylating agents on the chromatographic behavior of cysteine-containing peptides in reversed-phase chromatography

Xuehui Jiang[a,b], Dmitry Shamshurin[c,d], Vic Spicer[b,c], Oleg V. Krokhin[c,d,*]

[a] Department of Precision Instrument and Opto-Electronics Engineering, Tianjin University, Tianjin 300072, China
[b] Department of Physics and Astronomy, University of Manitoba, Winnipeg R3T 2N2, Canada
[c] Manitoba Centre for Proteomics and Systems Biology, University of Manitoba, 799 JBRC, 715 McDermot Avenue, Winnipeg R3E 3P4, Canada
[d] Department of Internal Medicine, University of Manitoba, 799 JBRC, 715 McDermot Avenue, Winnipeg R3E 3P4, Canada

ARTICLE INFO

Article history:
Received 10 July 2012
Accepted 12 December 2012
Available online 22 December 2012

Keywords:
Peptide reversed-phase HPLC
Chemically modified peptides
Peptide retention prediction
Cysteine alkylation

ABSTRACT

We investigate the influence of various alkylation chemistries on the reversed phase (RP) HPLC behavior of Cys-containing peptides under the most popular RP-HPLC conditions used in proteomics: C18 phases with trifluoroacetic acid (TFA) or formic acid (FA) as the ion pairing modifiers, and separation at pH 10. Akylating agents studied are iodoacetamide (IAM), iodoacetic acid (IAA), 4-vinylpyridine (4-VP), acrylamide (AA) and methyl methanethiosulfonate (MMTS). These were compared against the retention of identical peptides without alkylation, i.e. free cysteines. The intrinsic hydrophobicity values of the Cys residue under formic acid conditions for these modifications were found to increase in the following order: 4-VP < IAM < AA < IAA < free Cys < MMTS. The retention contribution of the positively charged 4-VP derivative is affected by the nature of the ion-pairing modifier; this is the most hydrophilic residue for formic acid based eluents, and second most hydrophilic behind IAM-alkylated Cys using TFA eluent. Switching to a basic condition dramatically decreases the retention of free cysteine and IAA-alkylated analytes due to the ionization of side-chains. The opposite effect is observed for 4-VP, which become neutral at basic pHs. The careful measurement of the hydrophobic contributions for these residues is vital to the development of accurate peptide retention prediction models; the incorporation of these modifications into our Sequence Specific Retention Calculator model is presented.

1. Introduction

The use of chemical reactions involving the alkylation of sulfhydryl groups has a long history [1,2]. The unique nucleophilic properties of the cysteine side chain make it a key element of many biochemical techniques involving protein modifications [3,4]. For example, Cys protection protocols are utilized in Edman degradation and the study of protein sequences by enzymatic hydrolysis [5,6]. Currently, the vast majority of proteomic experiments are performed bottom-up, i.e. proteins are reduced, alkylated, digested, and the resulting peptide mixtures are analyzed by liquid chromatography–mass spectrometry (LC–MS). Protein denaturation, accompanied by the breakage of disulfide bridges and the protection of free cysteines, is crucial for successful protein digestion [7]. Cysteine-containing peptides are detected in their alkylated forms. The basis for MS detection of derivatized peptides is well understood: the known chemical structures of alkylating agents introduce an easily calculated shift in peptide mass that can be measured by MS to high precision. However, knowing up-front the chromatographic behavior of a modified peptide has additional value. The LC component of the analysis has a much lower separation capacity (resolution) than MS, as well as more complex mechanisms driving its selectivity. In other words, the hydrophobicity of peptides cannot be calculated and measured with the same precision as their masses.

Understanding the mechanisms of peptide RP HPLC, and the prediction of peptide retention has been the pursued by chromatographers for over 30 years [8,9]. Originally these studies were directed at predicting and developing better separation protocols of peptidic compounds. With the arrival of proteomics, these studies were additionally motivated by the possible use of peptide retention time as an additional constraint in protein identification and characterization [10,11]. A number of new peptide retention prediction models have been developed in the past 8 years, significantly advancing our understanding of peptide RP separation mechanisms [12–17]. Studying the chromatographic behavior of peptides with chemical modifications was not a significant part of this process until recently. Early attempts were made to study

* Corresponding author at: Manitoba Centre for Proteomics and Systems Biology, 799 JBRC, 715 McDermot Avenue, Winnipeg, R3E 3P4, Canada. Tel.: +1 204 789 3283; fax: +1 204 480 1362.
E-mail address: krokhino@cc.umanitoba.ca (O.V. Krokhin).

1570-0232/$ – see front matter © 2013 Published by Elsevier B.V.
http://dx.doi.org/10.1016/j.jchromb.2012.12.010

phosphorylated peptides for their inclusion into retention prediction protocols, as phosphorlyation is one of the most biologically relevant post-translational modifications (PTMs) [18]. Our group began systematic studies of the RP HPLC of modified peptides with N-terminal cyclization at Gln and carbamidomethyl-Cys [19]. The final goal of these (and similar ongoing) studies is to expand applicability of our Sequence Specific Retention Calculator (SSRCalc) model.

Early retention prediction models were developed for unprotected cysteine residues, as most of those studies were based on the analysis of custom-designed synthetic peptides using UV detection [20]. In the proteomic era, the vast majority of peptide retention prediction models were developed using real proteomics samples, where alkylation is mandatory. Iodoacetamide (IAM) alkylation is the most popular cysteine-blocking agent; most of the models were optimized using IAM. The work by Petritis et al. [14] was the exception, working with unprotected peptides. The choice of alkylating agents for this study was based on their popularity and importance for shotgun proteomic applications. Iodoacetamide, iodoacetic acid, 4-vinylpyridine are by far the most used alkylating agents [7]. Acrylamide is important because of the partial alkylation of Cys residues during SDS-page separation protocols and subsequent in-gel protein digestion [7,21]: residual amounts of acrylamide in polyacrylamide gels gives rise to a mix of alkylated versions of Cys-containing peptides. Methyl methanethiosulfonate is widely used due to the reversible character of labeling [22]. It is also the recommended alkylating agent for iTRAQ quantitative analysis protocols – a significant number of quantitative proteomic studies are performed using MMTS [23]. We choose to compare all of these alkylation chemistries against unprotected cysteines – the most basic alkylation state for peptide synthesis, purification and analysis. The chromatographic conditions for our study cover the most popular peptide-oriented proteomic applications: C18 100 Å sorbents with trifluoroacetic (used in off-line LC–MALDI MS [24]), formic acid (used in LC–ESI MS) as ion-pairing modifiers, and separation at basic pH10 (used as the first dimension in 2D-LC MS protocols [25]). Understanding the chromatographic properties of cysteine-containing peptides under various alkylation chemistries and their inclusion in retention prediction modeling would significantly benefit both the chromatographic and proteomic research communities.

2. Materials and methods

2.1. Chemicals

Deionized (18 MΩ) water and HPLC-grade acetonitrile were used in eluent preparation. All chemicals were sourced from Sigma Aldrich (St.-Louis, MO) unless otherwise noted. Sequencing-grade modified trypsin (Promega, Madison, WI) was used for protein digestion. A number of commercially available proteins were used to generate mixtures of peptides for off-line LC–MALDI-MS or LC–ESI-MS/MS analysis and assembly of the peptide retention datasets: human albumin and lactoferrin, bovine apotransferrin and albumin, chicken conalbumin, sheep and porcine albumins.

2.2. Sample preparation

The stock solutions of proteins (1 mg/ml) were prepared in 100 mM NH_4HCO_3 buffer and mixed in equal amounts. Sample preparation steps included reduction (10 mM dithiothreitol (DTT), 30 min, 57 °C), alkylation, dialysis (6 h using a 7 kDa molecular weight cut-off filter (Pierce, Rockford, IL) against 100 mM NH_4HCO_3), and finally trypsin digestion (1:50 enzyme/substrate weight ratio for 12 h at 37 C). Fig. 1 shows the chemical structure of the alkylating agents. Cysteine alkylation procedures were as follows:

(1) *Iodoacetamide (IAM), Iodoacetic acid (IAA), acrylamide (AA), methyl methanethiosulfonate (MMTS).* 10 μl of an alkylating reagent stock solution (500 mM in 100 mM NH_4HCO_3 buffer) was added to 100 μl of the reduced sample (to yield a 5:1 molar ratio of alkylating agent to DTT). The resulting mixture was incubated in the dark for 1 h at room temperature.
(2) *4-Vinylpyridine (4-VP).* 10 μl of a 4-vinylpridine solution (500 mM in methanol) was added to 100 μl of the reduced sample. The resulting mixture was incubated in dark for 1 h at room temperature.
(3) *Free Cys.* Alkylation step was omitted in this case. Small amount of DTT was added to the samples prior to the injection into the chromatographic system to reduce disulfide bonds formed during the digestion.

Peptide mixtures were diluted prior to LC–MS analysis to provide ~2–3 pmol and ~100–200 fmol injections of all components for LC–MALDI MS and LC–ESI MS analyses, respectively. Samples were also supplemented with the addition of our P1–P6 6-peptide standard mixture [26] to facilitate data alignment.

2.3. HPLC instrumentation

A micro-Agilent 1100 Series system (Agilent Technologies, Wilmington, DE) was used for the off-line LC–MALDI-MS experiments [24] with a direct sample injection (loop size 10 μl). Digests were fractionated on a 300 μm × 150 mm PepMap100 column (Dionex, Sunnyvale, CA) with 0.1% trifluoroacetic acids as the ion-pairing modifier, and on a 150 μm × 150 mm XTerra column (Waters, Milford, MA) with 20 mM ammonium formate, at pH 10 the for pH 10 separation experiments. A linear water – acetonitrile gradient of 1% per minute was used for all cases, starting at 1% of organic solvent. The column effluent (3 μL/min) was mixed on-line with a 2,5-dihydroxybenzoic MALDI matrix solution (150 mg/ml in water:acetonitrile 1:1; 0.5 μL/min), and robotically deposited on MALDI target at 36 s intervals. These spots were air-dried and subjected to MALDI-MS analysis.

A split-less nanoflow Tempo LC system (Eksigent, Dublin, CA) with a 20 μL sample injection via a 300 μm × 5 mm PepMap100 pre-column, and a 100 μm × 150 mm analytical column packed with 5 μm Luna C18(2) (Phenomenex, Torrance, CA) was used for the LC–ESI-MS/MS analyses. Both eluents A (0.5% acetonitrile in water) and B (98% acetonitrile) contained 0.1% formic acid as the ion-pairing modifier. Digest were fractionated using a linear gradient of 1% acetonitrile per minute starting from 0% B with 500 ml/min flow rate.

2.4. Mass spectrometry, peptide identification, retention time assignment

MALDI spectra of chromatographic fractions were acquired using the Manitoba/Sciex prototype MALDI quadrupole/TOF mass spectrometer with 10 ppm mass accuracy in both MS and MS/MS modes [27]. Peptides were first tentatively assigned using our in-house LC–MALDI-MS peptide mass-fingerprinting engine, which uses MAss and Retention Time (sMART) as identification constraints [28]. Peptide identities have been confirmed by MS/MS measurements. The fraction number was used as a measure of peptide retention time. If a peptide peak was contained within a single spot fraction, the peak was assigned a retention time equal to the fraction number. However, if the peak MS signal was split across two (or sometimes three) consecutive fractions, the retention time assigned based on the intensity-weighted average of the two most

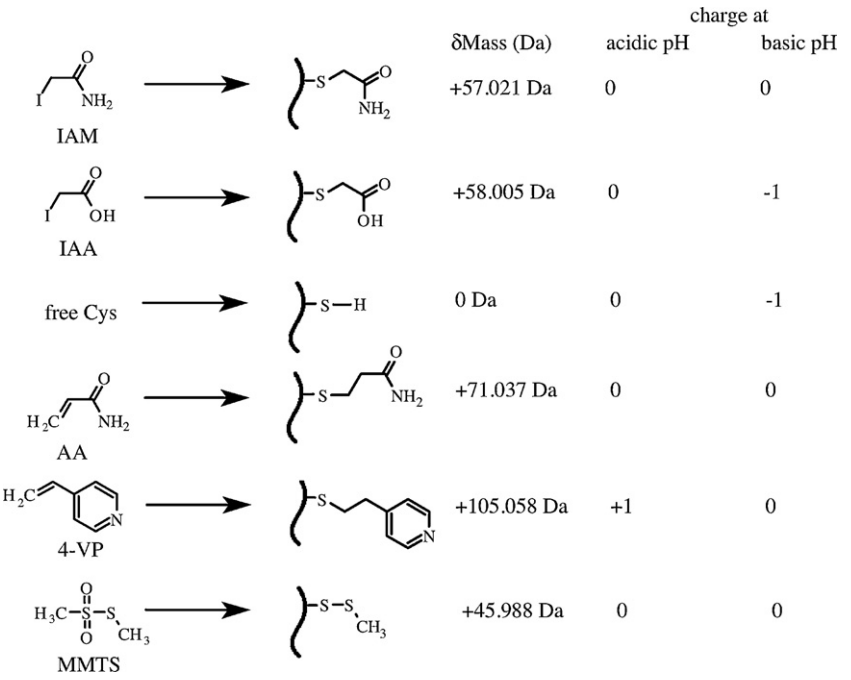

Fig. 1. Chemical structures of modified cysteine side chains and alkylating agents under investigation in this study.

intense fractions. Retention times were determined multiplying fraction number by 0.6 min – the size of one fraction.

A QStar Elite mass spectrometer (ABSciex, Toronto, ON) was used in standard MS/MS data-dependent acquisition mode. One second of survey MS spectra were collected (m/z 300–1500) followed by 3 MS/MS measurements on the most intense parent ions (80 counts/s threshold, +2/+3/+4 charge states, m/z range 100–1500 for MS/MS). Previously interrogated parent ions were excluded from repetitive MS/MS acquisition over 60 s elution window. Raw spectra files were treated using the packaged Mascot.dll script (Analyst QS2.0) to create text files in the Mascot Generic File format. Protein/peptide identification was performed using the X!Tandem (GPM) search engine. Standard QTOF settings were used for the search: 100 ppm and 200 ppm mass tolerance for parent and fragment ions, respectively. The retention times of identified peptides were assigned manually as peak maxima on extracted ion chromatograms.

LC–MALDI MS and LC–ESI MS (MS/MS) differ in how they couple the chromatographic system to the mass spectrometer. LC–MALDI MS is an inherently off-line procedure, allowing independent operation of both LC and MS components and the comprehensive interrogation of particular fraction if needed, but it is more time consuming compared to LC–ESI. In both instances we used the identical gradient slope: 1% acetonitrile per minute for 40 min, followed by washing and equilibration steps (20 min combined). ESI MS acquisition occurs on-line, concurrent with the (1 h) LC separation. MALDI MS acquisition required at least one minute per fraction (over 60 fractions) when operated in manual mode, effectively doubling the overall analysis time. Subsequent MALDI MS/MS for peptide sequence confirmation extends this analysis time even further. Despite the obvious productivity advantages of LC–ESI, the inherent incompatibility of both TFA-based and high-pH eluents with the electrospray ionization mechanism required the use of MALDI MS.

2.5. Calculations and programming

The core SSRCalc program was implemented in Perl. The output of retention prediction was presented as t_R vs. *HI* graphs, where *HI* (hydrophobicity index) represents calculated acetonitrile percentage, required for elution of a peptide under experimental conditions [26]. Retention data sets contained identical non-Cys peptides and Cys-containing species with a particular modification. The optimization consisted of adjusting R_c (retention coefficient) values for Cys residue in different alkylation states to provide best possible as t_R vs. *HI* correlation, while also holding the slope and intercept values identical (or close) to that from the non-Cys species.

3. Results and discussion

3.1. Assignment of retention times of peptides under different alkylation chemistries

Fig. 2 shows the MALDI MS spectra of fractions containing the bovine apo-transferrin peptide KPVTDAENCHLAR carrying different Cys-blocking groups and the non-modified Cys residue. Known mass difference (Fig. 1) and accurate mass measurement (±10 ppm) allow for confident assignment of the different species and their retention times (fraction numbers). As noted previously, most peptide peaks are distributed between two or more fractions and the fraction number is the intensity weighted average over the two most intense fractions. Conversely, the online coupling of RPLC and ESI MS provides a continuous trace of extracted ion chromatograms, yielding a more accurate assignment of peptide retention times. Following peptide retention time assignments, all chromatogram were aligned using an in-house peptide retention standard as described in [26]. The combined retention data sets for all three chromatographic conditions across the six different alkylation chemistries are provided in Appendix A. These sets contain between 244 and 340 peptides, with about half of them carrying Cys residue(s). It should be noted that for real tryptic digests of complex protein mixtures, typically only 20–25% of the peptides contain a cysteine. This difference is observed because we selected Cys-rich proteins for this study.

3.2. Factors affecting hydrophobicity of modified side chain of Cys residue

The alkylation agents studied can be arranged in order of increasing hydrophobicity of the Cys side chain:

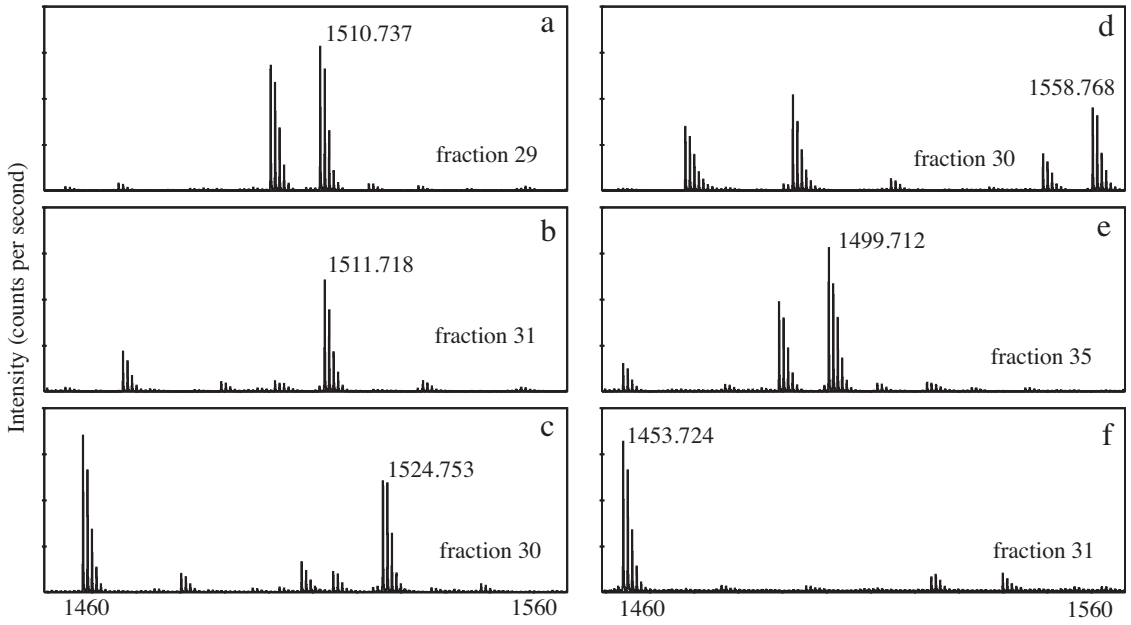

Fig. 2. HPLC/MALDI-MS spectra of KPVTDAENCHLAR peptide from bovine transferrin and its cysteine alkylated products with different reagents. Shown for TFA-based eluent system. (a) IAM alkylation, fraction 29; (b) IAA alkylation, fraction 31; (c) AA alkylation, fraction 30; (d) 4-VP alkylation, fraction 30; (e) MMTS alkylation, fraction 35; (f) free Cys, fraction 31.

TFA conditions: IAM < 4-VP < AA < IAA < Cys < MMTS
FA conditions: 4-VP < IAM < AA < IAA < Cys < MMTS
pH 10 conditions: IAA < Cys < IAM < AA < 4-VP = MMTS

The alkylating agents' chemical structures (Fig. 1) can inform this retention ordering. For example, peptides modified by AA retained higher compared to IAM across all three eluent conditions due to the presence of an extra methylene group in the structure of AA. The structure of MMTS-modified Cys is very similar to methionine – the 5th most hydrophobic residue among all naturally occurring amino acids. Indeed, experimental results show that MMTS modified peptides possesses highest retention across the three eluent conditions studied. But the relative retention of iodoacetamide and iodoacetic acid derivatives is harder to assess based solely on their structure. It is known however, that acidic analogs in Asp-Asn and Glu-Gln pairs exhibit higher retention coefficients at acidic conditions [13]. The same situation is observed in the case of IAM and IAA acetylated Cys – the more acidic iodoacetic acid derivative shows a higher hydrophobicity. Ionization of Asp and Glu at higher pH leads to a dramatic decrease in their retention coefficients [25]. The same is true for the IAM and IAA pair: while for acidic eluents the IAA analogs retain stronger, at pH 10 it becomes the most hydrophilic from all studied modifications.

While analyzing RPLC retention contributions, particular attention should be paid to the charge state of the side chains at the pH of the eluent. Residues carrying charged functional groups tend to be hydrophilic. Fig. 1 shows the side chain charge at acidic and basic pHs. Three out of six groups will change their charge state upon switching pH, providing a significant alteration in retention properties: IAA, 4-VP and free Cys. Similar to the IAA case discussed above, free Cys is the second-most hydrophobic under TFA and FA conditions, and the second most hydrophilic at pH 10 (behind IAA derivative). Interestingly, the order IAA < Cys is the same under both basic and acidic pH. 4-VP is hydrophobic at basic pH but hydrophilic at acidic pH, due to the protonation of nitrogen in the pyridine ring.

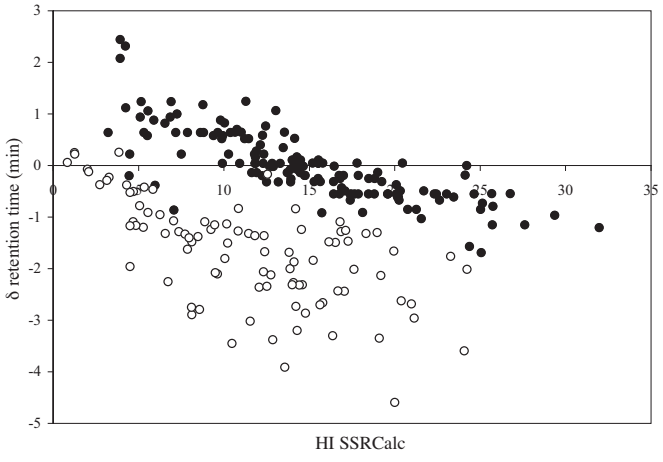

Fig. 3. Retention time difference between 4-VP and IAM alkylated peptides ($\delta = t_{R(4\text{-VP})} - t_{R(IAM)}$, min) separated under TFA (solid circles) and FA (empty circles) conditions. Hydrophobicty Index values were calculated for IAM-alkylated members of respective pairs.

3.3. *The effect of ion-pairing modifier hydrophobicity on intrinsic hydrophobicity values of charged residues*

The retention ordering for TFA and FA conditions are similar. This is not surprising, as these two eluent systems are at similar pH values. The only exception is the 4-VP – IAM pair: 4-VP is found to be the most hydrophilic under FA conditions. This difference can be explained based on the hydrophobic properties of ion-pairing agents. Indeed, the more hydrophilic formate anion decreases the retention of the positively charged 4-VP group to a greater extent compared to the less hydrophilic trifluoroacetate counter-anion, shown in Fig. 3. This effect of is well documented in peptide RP-HPLC literature [29].

Detailed observations of the relative retention of pairs of 4-VP – IAM modified peptides reveals a more complex picture. In TFA based systems, the retention of 4-VP modified peptides could be either higher or lower compared to its IAM counterpart. Fig. 3 shows the dependence of retention time difference between 4-VP and IAM modified species ($t_{R(4\text{-VP})} - t_{R(IAM)}$) from the overall peptide hydrophobicity. For TFA, switching from IAM

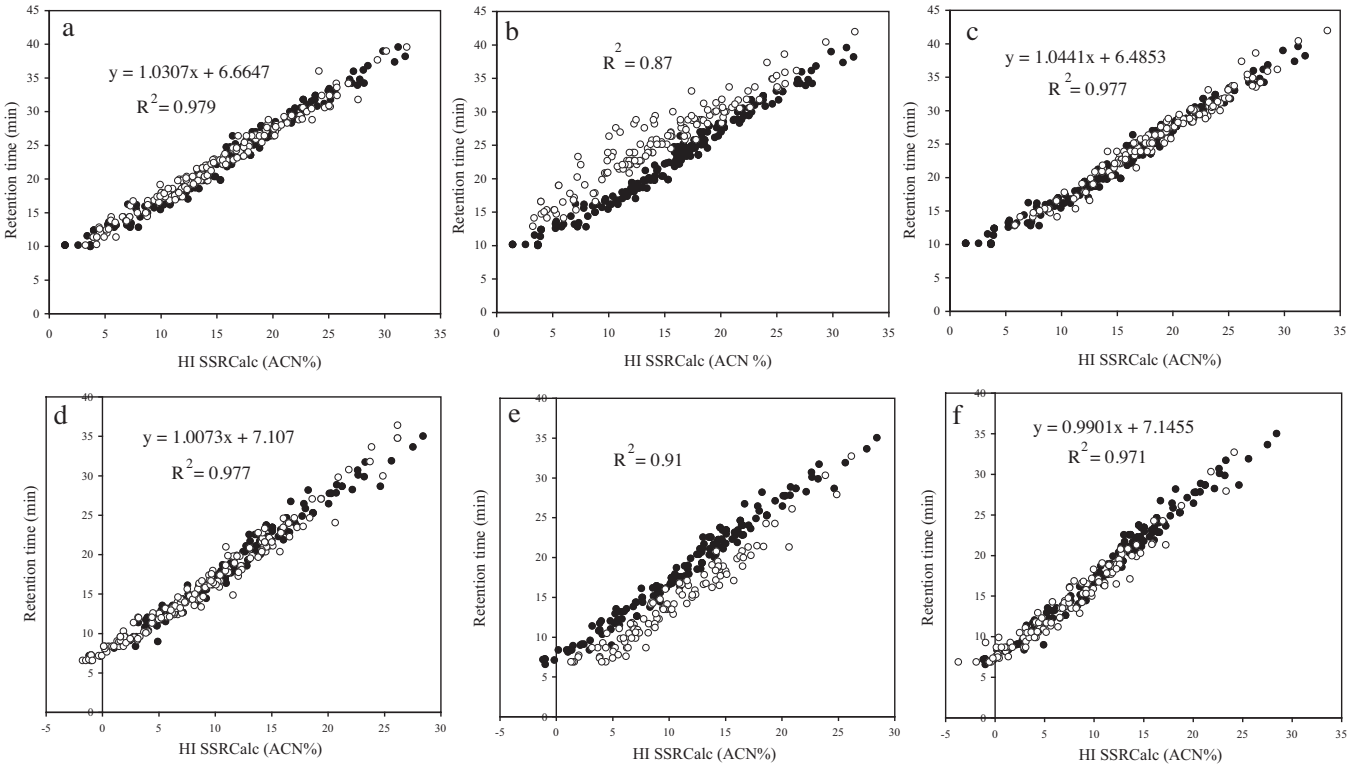

Fig. 4. Optimization of SSRCalc models for differentially alkylated Cys-containing peptides. (a–c) TFA – based eluent; (a) SSRCalc retention prediction for IAM (standard alkylation); (b) and (c) retention prediction for the dataset containing MMTS alkylated peptides before and after optimization; (d–f) pH 10 eluent conditions; (d) SSRCalc retention prediction for IAM (standard alkylation); (e) and (f) retention prediction for the dataset containing IAA alkylated peptides before and after optimization; empty circles – Cys containing peptides; solid circles – all other peptides.

to the "more hydrophilic" charged 4-VP decreases the retention for more hydrophobic species, as well as the reverse: it increases for hydrophilic species. The behavior of the same peptides in FA-based systems is similar, except the retention shift is more negative. Only a few the most hydrophilic peptides exhibit higher retention of the 4-VP modified species as shown in Fig. 3.

Recently we observed similar effects when attempting to derive intrinsic hydrophobicity values for naturally occurring Lys, Arg and His – all positively charged at acidic pH [30]. We concluded that assignment of hydrophobicity for charged residues is ambiguous and cannot be done without taking into consideration the overall peptide hydrophobicity. The same is true for 4-VP derivatives in both TFA and FA – based eluents. Therefore it is important to remember that the value of the retention coefficient for charged residues is based on a model optimization over a population of peptides. Thus, the retention coefficients of 4-VP and IAM modified cysteines for TFA based eluent were found to be very close – result of averaging the contribution for all species in the data set. But the experimental data in Fig. 3 shows that difference in retention time for these pairs varies within ±2 min (2% acetonitrile).

3.4. SSRCalc retention time prediction for Cys-containing peptides with different alkylation chemistry

Fig. 4(a) shows the SSRCalc retention prediction applied to TFA – based separation for a typical set of tryptic peptides alkylated with iodoacetamide. Since we used HI units (acetonitrile percentage) to express peptides' hydrophobicity with a 1%-per-minute acetonitrile gradient, this plot has the slope very close to 1. Cys-containing peptides (empty circles) show the same prediction accuracy as the rest of the analytes (solid circles). Application of the same (unmodified) version of the model to the set with MMTS alkylation shows a significantly lower correlation due to the increased retention of Cys-containing peptides (Fig. 4(b).

The retention coefficient for each Cys alkylation state was optimized with two objectives: (1) to maximize the Pearson's correlation coefficient between the calculated HI values from the SSRCalc algorithm and the observed peptide dataset's retention time values; (2) to minimize the deviation of the resulting correlation's slope and intercept values from their unmodified Cys counterparts. This was achieved using the manual implementation of a golden section search, where a range of upper and lower bounds on values of the retention coefficient (for each modification) were iteratively narrowed based on the R^2-value returned from the respective SSRCalc algorithm.

In case of MMTS shown in Fig. 4(b), the retention coefficient for the modified cysteine was altered to provide the best possible correlation for the combined set of Cys-containing peptides and all other peptides (Fig. 4(c). The resulting correlation of t_R vs. HI yields similar slope and intercept values to that in Fig. 4(a) because the subset of non-Cys peptides is identical across both datasets. Not surprisingly, the value of retention coefficient for MMTS-Cys (Table 1) was found to be much higher than that for IAM.

Similar picture was observed for the optimization of the IAA model under pH 10 conditions (Fig. 4(d–f)). In this case, however, the Cys containing peptides were below the trend line prior to optimization (Fig. 4(e)). This is a consequence of the significantly lower hydrophobicity of the side chain with negatively charged carboxy group at pH 10. Model optimization allowed for a significant improvement of observed correlation for Cys-containing peptides alkylated with iodoacetic acid (Fig. 4(f)). At pH 10, IAA-Cys is the most hydrophilic of all modifications studied (Table 1).

Table 1 compares the optimized retention coefficients for all naturally occurring amino acids, including the 6 different alkylation chemistries of Cys. Most of these cysteine modifications are

Table 1
Retention coefficients (R_c) for the optimized SSRCalc models.

Residue	100 Å – TFA	100 Å – FA	pH 10
W	13.12	13.45	12.27
F	11.34	11.70	10.19
L	9.44	10.19	8.74
I	7.86	8.95	7.47
M	6.57	6.65	5.67
V	4.86	5.64	4.86
Y	5.40	5.86	4.77
P	1.62	2.42	1.85
A	1.11	1.49	1.57
E	1.08	1.95	−4.94
T	0.48	1.12	1.06
D	−0.22	1.06	−5.41
Q	−0.53	0.21	0.30
S	−0.33	0.25	0.61
G	−0.35	0.02	0.17
R	−2.58	−3.83	3.56
N	−1.44	−0.74	0.04
H	−3.04	−4.50	0.66
K	−3.53	−4.48	2.80
C-IAM	0.04	0.40	0.34
C-AA	0.85	0.60	0.52
C-IAA	2.15	3.15	−5.45
C-4VP	0.24	−3.20	7.74
C-free	3.85	3.90	−0.54
C-MMTS	8.56	10.51	7.74

moderately hydrophilic (positioned in the middle of hydrophobicity scale of naturally occurring residues). MMTS-modified cysteine has a hydrophobicity value between Ile and Leu, making it the 4th most lipophilic residue behind Trp, Phe and Leu. Note that 4-VP – modified cysteine has the same retention coefficient (hydrophobicity) as MMTS only at pH 10. The same residue is 4th most hydrophilic under formic acid conditions after only Lys, His and Arg. Free cysteine exhibit weak hydrophobic properties under acidic conditions, slightly below that of Val.

It should be noted that the optimized retention coefficient values in Table 1 and the retention ordering shown in Section 3.2 represent the average values for each modification: these orderings could be slightly different for specific peptides. For example, DVFSECCQAADK derivatives under formic acid conditions follow the general ordering shown in Table 1: 4-VP < IAM < AA < IAA < Cys < MMTS. Conversely, the retention ordering for CLAENAGDVAFVK is: 4-VP < IAM < AA < Cys < IAA < MMTS. The higher retention of the IAA analog compared to its free cysteine counterpart for the latter peptide could be the effect of acidic modifying moiety on basicity of the N-terminal amino group. Decreasing the basicity leads to lowering the ion-pair formation at the positively charged N-terminus, which involves interactions with hydrophilic formate counter-ions. This decrease in ion-pairing formation increases the hydrophobicity of the IAA modified peptide. The same CLAENAGDVAFVK peptide under TFA conditions shows the conventional ordering of IAA < Cys due to the higher hydrophobicity of the ion-pairing modifier. Overall, higher hydrophilicity of the formate anion compared to trifuoroacetate leads to both decreased retention for basic 4-VP analogs, and increased retention for acidic IAA derivatives.

These examples illustrate that developing a comprehensive retention prediction model will require taking into account all possible sequence specific factors, and building an accurate framework of their interactions. Currently SSRCalc models do not take into account the interactions of N-terminal amino and C-terminal carboxy groups depending on their neighboring amino acids; these examples illustrate the necessity for further developments in retention prediction modeling.

4. Conclusion

The development of peptide retention prediction models continues through both improving the prediction accuracy by further studies of retention mechanisms, and the inclusion of various chemical/post-translational modifications in the modeling process. The latter extends the model's applicability to a variety of modifications encountered in proteomic experiments. We investigated the influence of various alkylation chemistries on the reversed phase (RP) HPLC behavior of Cys-containing peptides, and modified our SSRCalc retention prediction model to accommodate these changes. The resulting values of retention coefficients for various alkylation states of Cys illustrate the hydrophobic properties of the Cys side chain under different eluent conditions we used: trifluoroacetic acid and formic acids as ion pairing modifiers, and separation at pH 10. We find that for acidic TFA eluent conditions the following order is generally observed: IAM < 4-VP < AA < IAA < Cys < MMTS. Switching to formic acid mostly affects the order of IAM 4-VP pair – due to higher hydrophilicity of the formate counter ion positively charged 4-VP side chain exhibit more hydrophilic properties. Another difference is higher value for iodoacetic acid's retention coefficient, which leads to the alternative retention order Cys < IAA for some peptides. Both these effects could be explained from the point of view of ion-pairing formation, which involve more hydrophilic formate counter ion. Acidic character of IAA modified side chain decrease ion-pair formation resulting in overall increase of peptide hydrophobicity. The effect of basic substitution such as 4-VP is opposite.

Varying pH of the eluent profoundly affects the hydrophobicity values of ionogenic resides, which alter their charge status. Thus, Cys alkylated with iodoacetic acid becomes the most hydrophilic residue at pH 10 due to dissociation of the carboxy group. Conversely, 4-VP looses its proton at pH 10 and becomes very hydrophobic.

The methodology developed for this study serves as a reference for further directed investigations of chromatographic behavior of peptides carrying PTMs. The core steps include generating a representative peptide library, carefully applying the modifications, identifying peptide sequences and their retention times to high confidence and accuracy, alignment of retention times across the various bottom-up LC-MS experiments by using peptide retention standards, and the automated or human-directed re-optimization of retention parameters. Appreciating the chemical and separation-level implications of the PTMs themselves will allow us to pursue wide variety of biologically relevant PTMs which impact disease processes and functional regulations.

Acknowledgements

This work was supported in part by the Natural Sciences and Engineering Research Council of Canada (O.V.K.) and the China Scholarship Council (X.J.).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jchromb.2012.12.010.

References

[1] F. Dickens, Biochem. J. 27 (1933) 1141.
[2] C.V. Smythe, J. Biol. Chem. 114 (1936) 601.
[3] G.L. Kenyon, T.W. Bruice, Methods Enzymol. 47 (1977) 407.
[4] M.W. Crankshaw, G.A. Grant, Modification of cysteine, in: Current Protocols in Protein Science, Wiley Interscience, Hoboken, NJ, 2001, Chapter 15, Unit 15.1.

[5] D. Brune, Anal. Biochem. 207 (1992) 285.
[6] A.M. Crestfield, S. Moore, W.H. Stein, J. Biol. Chem. 238 (1963) 622.
[7] S. Sechi, B.T. Chait, Anal. Chem. 70 (1998) 5150.
[8] J.L. Meek, Proc. Natl. Acad. Sci. U. S. A. 77 (1980) 1632.
[9] C.A. Browne, H.P. Bennett, S. Solomon, Anal. Biochem. 124 (1982) 201.
[10] E.F. Strittmatter, L.J. Kangas, K. Petritis, H.M. Mottaz, G.A. Anderson, Y. Shen, J.M. Jacobs, D.G. Camp 2nd, R.D. Smith, J. Proteome Res. 3 (2004) 760.
[11] M. Palmblad, M. Ramstrom, K.E. Markides, P. Hakansson, J. Bergquist, Anal. Chem. 74 (2002) 5826.
[12] O.V. Krokhin, R. Craig, V. Spicer, W.E. Ens, K.G. Standing, R.C. Beavis, J.A. Wilkins, Mol. Cell. Proteomics 3 (2004) 908.
[13] O.V. Krokhin, Anal. Chem. 78 (2006) 7785.
[14] K. Petritis, L.J. Kangas, B. Yan, M.E. Monroe, E.F. Strittmatter, W.J. Qian, J.N. Adkins, R.J. Moore, Y. Xu, M.S. Lipton, D.G. Camp 2nd, R.D. Smith, Anal. Chem. 78 (2006) 5026.
[15] A.V. Gorshkov, I.A. Tarasova, V.V. Evreinov, M.M. Savitski, M.L. Nielsen, R.A. Zubarev, M.V. Gorshkov, Anal. Chem. 78 (2006) 7770.
[16] L. Moruz, D. Tomazela, L. Kall, J. Proteome Res. 9 (2010) 5209.
[17] M. Gilar, H. Xie, A. Jaworski, Anal. Chem. 82 (2010) 265.
[18] J. Kim, K. Petritis, Y. Shen, D.G. Camp 2nd, R.J. Moore, R.D. Smith, J. Chromatogr. A 1172 (2007) 9.
[19] J. Reimer, D. Shamshurin, M. Harder, A. Yamchuk, V. Spicer, O.V. Krokhin, J. Chromatogr. A 1218 (2011) 5101.
[20] D. Guo, C.T. Mant, A.K. Taneja, J.M.R. Parker, R.S. Hodges, J. Chromatogr. 359 (1986) 499.
[21] M. Chiari, P.G. Righetti, A. Negri, F. Ceciliani, S. Ronchi, Electrophoresis 13 (1992) 882.
[22] D.J. Smith, E.T. Maggio, G.L. Kenyon, Biochemistry 14 (1975) 766.
[23] P.L. Ross, Y.N. Huang, J.N. Marchese, B. Williamson, K. Parker, S. Hattan, N. Khainovski, S. Pillai, S. Dey, S. Daniels, S. Purkayastha, P. Juhasz, S. Martin, M. Bartlet-Jones, F. He, A. Jacobson, D.J. Pappin, Mol. Cell. Proteomics 3 (2004) 1154.
[24] O.V. Krokhin, W. Ens, K.G. Standing, J. Biomol. Tech. 16 (2005) 429.
[25] R.C. Dwivedi, V. Spicer, M. Harder, M. Antonovici, W. Ens, K.G. Standing, J.A. Wilkins, O.V. Krokhin, Anal. Chem. 80 (2008) 7036.
[26] O.V. Krokhin, V. Spicer, Anal. Chem. 81 (2009) 9522.
[27] A.V. Loboda, A.N. Krutchinsky, M. Bromirski, W. Ens, K.G. Standing, Rapid Commun. Mass Spectrom. 14 (2000) 1047.
[28] O.V. Krokhin, V. Spicer, K.G. Standing, J.A. Wilkins, W. Ens, 54th ASMS Conference on Mass Spectrometry and Allied topics, Seattle, USA, poster, 2006.
[29] D.C. Guo, C.T. Mant, R.S. Hodges, J. Chromatogr. 386 (1987) 205.
[30] D. Shamshurin, V. Spicer, O.V. Krokhin, J. Chromatogr. A 1218 (2011) 6348.