BIOTECHNOLOGY - DERIVED ARTICLES
PEPTIDE MAPPING

Revision 1, Stage 4

INTRODUCTION

Peptide mapping is an analytical method used to characterize the primary structure of proteins, in this context those obtained through recombinant DNA technology. Peptide mapping involves the chemical or enzymatic treatment of a protein to selectively cleave amide bonds between amino acid residues to yield a predictable set of peptides. The peptides produced are subjected to analytical separation, detection, and identification, providing information on the amino acid sequence and any chemical modification of the peptides. Once detailed chemical structural information for peptides in the peptide map has been determined, this information can be extrapolated back to provide chemical structural information for the protein from which the peptide map was derived. Peptide mapping is a comparative procedure because the information obtained, compared to a product reference standard similarly treated, confirms the primary structure of the protein, is capable of detecting whether alterations in structure have occurred, and demonstrates process consistency and genetic stability.

Within the scope of the analysis of biotechnology-derived protein pharmaceuticals, there are several intended uses of peptide mapping. This chapter will focus on the use of peptide mapping for the determination of the primary structure of a protein to establish identity of a new product or confirm identity of existing products This is achieved by comparison of the peptide map obtained with the test substance to that obtained with a reference substance similarly treated.

GENERAL CONSIDERATIONS FOR THE DEVELOPMENT OF A PEPTIDE MAPPING PROCEDURE

Each protein to be mapped presents unique characteristics that must be well understood so that the scientific and analytical approaches permit validated development of a peptide map that provides sufficient specificity. In addition, the design of peptide mapping procedures may differ according to the goal of the procedure. For protein identity, a peptide map needs high specificity. Complete or nearly complete sequence coverage is important, because there may be no prior knowledge of the alterations to the protein. There are general considerations that apply to all peptide mapping applications.

Four major steps are typically necessary for the development of a peptide mapping procedure: sample pre-treatment (e.g., desalting, reduction, alkylation); selective cleavage of the amide bonds; peptide separation; and detection of the peptides. This chapter provides detailed guidance in the application of peptide mapping approaches.

43 The flow chart in Figure 1 outlines the steps and decisions involved in
44 developing a peptide mapping procedure. In general, development of a peptide
45 map procedure starts with knowledge of the amino acid sequence of the protein
46 product. Because the chemical or enzymatic proteolysis methods used in
47 peptide mapping have specificity towards amide bonds prior to or after specific
48 amino acids, knowledge of the amino acid sequence can be used to predict the
49 set of peptides that will be formed. Computational tools are available to predict
50 the peptides generated by a chemical reagent capable of cleaving a sequence
51 specific peptide bond or a proteolytic enzyme, i.e., for performing an *in silico*
52 digestion. For each peptide mapping application, *in silico* digestion is first
53 performed to both "design" the set of hypothetical proteolytic peptides that, once
54 characterized, will achieve the goal of the peptide map and determine which
55 proteolytic reagent to use. After performing the *in silico* digestion experiment, the
56 next step in developing a peptide mapping procedure is to perform the selected
57 proteolytic process followed by analytical characterization of the peptides
58 generated. The techniques used for analytical characterization can vary, but
59 typically involve a method to first separate the peptides (such as reversed-phase
60 liquid chromatography, (RPLC)) and then to detect and identify separated
61 peptides (such as UV detection or mass spectroscopy (MS) or Edman
62 sequencing of the isolated peptides). The analytical results are then compared
63 to those predicted by *in silico* digestion and optimization of the sample
64 processing, peptide separation, or detection can be performed iteratively or by
65 using Design of Experiments (DOE) if the experimental results do not sufficiently
66 match the expected results.

67

68 **EXPERIMENTAL ASPECTS OF PEPTIDE MAPPING**

69 Peptide mapping is not a general method, but involves developing a
70 specific map for each unique protein. Although the technology is evolving rapidly,
71 there are certain methods that are well established. Variations of these methods
72 will be indicated, when appropriate, in specific monographs.

73 **Computational Tools for Peptide Mapping**

74 There are a variety of computational tools available to assist in the
75 experimental design of a peptide map. To use one of these computational tools
76 for *in silico* peptide mapping, one starts with the amino acid sequence of the test
77 protein and selects a cleavage agent to be used in the generation of the peptide
78 map. Next, a list of potential peptides is generated for the selected cleavage
79 agent. Depending on the sophistication of the computational tool used, the list of
80 potential peptides can also include characteristics about each peptide including
81 molecular mass and charge state masses for MS analysis, calculated
82 hydrophobicities, and calculated retention factors for chromatographic analysis.
83 By examining the list of potential peptides for a variety of cleavage agents, one
84 can decide which cleavage agent (or combination of cleavage agents) may best
85 match the peptide mapping goals. Computational tools for peptide mapping are
86 often included as a component of protein sequence databases found on the

87 internet or as part of the protein/peptide data analysis software that is produced
88 by manufacturers of LC-MS instrumentation.

89

90 Pretreatment of Sample

91 Isolation and purification may be necessary for analysis of bulk drugs,
92 dosage forms or reference substances containing interfering excipients and
93 carrier proteins. The procedures and respective system suitability requirements
94 for such procedures are specified in the individual monographs.

95 The treatment of proteins with chaotropic agents (e.g., guanidinium
96 chloride, urea) and surfactants (e.g., SDS) can be used to unfold the protein prior
97 to digestion. It is often necessary to reduce and alkylate the disulfide bonds prior
98 to digestion in order to allow the enzyme to have full access to cleavage sites;
99 however, the cysteine to cysteine linkage information is then lost. Common
100 reagents for disulfide reduction include dithiothreitol and trialkylphosphine
101 compounds such as tris(2-carboxyethyl)phosphine. Reagents for alkylating
102 reduced cysteines include iodoacetamide, iodoacetic acid, and 4-vinylpyridine.

103

104 Digestion

105 Selective Cleavage of Peptide Bonds

106 The selection of the approach used for the cleavage of peptide bonds is
107 protein dependent. This selection process involves determination of the type of
108 cleavage to be employed—enzymatic or chemical—and the type of cleavage
109 agent within the chosen category. Several cleavage agents and their specificity
110 are shown in *Table 1*. This list is not all-inclusive and while only a few of the
111 methods are widely utilized, there may be specific reasons for using other
112 methods or combinations of methods.

113

114 **Table 1. Examples of Cleavage Agents**

Type	Agent	Specificity
Enzymatic	Trypsin, EC 3.4.21.4	*C*-terminal side of Arg and Lys
	Chymotrypsin, EC 3.4.21.1	*C*-terminal side of hydrophobic residues (e.g., Leu, Met, Ala, aromatics)
	Pepsin A (Pepsin), EC 3.4.23.1	Low specificity digest
	Lysyl endopeptidase (Lys-C endopeptidase), EC 3.4.21.50	*C*-terminal side of Lys
	Glutamyl endopeptidase (Glu-C endoproteinase; V8 protease);	*C*-terminal side of Glu and Asp

Type	Agent	Specificity
	(from *S. aureus* strain V8), EC 3.4.21.19	
	Peptidyl-Asp metalloendopeptidase (Asp-N endoproteinase), EC 3.4.24.33	*N*-terminal side of Asp
	Clostripain (Arg-C endopeptidase), EC 3.4.22.8	*C*-terminal side of Arg
Chemical	Cyanogen bromide	*C*-terminal side of Met
	2-Nitro-5-thiocyanobenzoic acid	*N*-terminal side of Cys
	O-Iodosobenzoic acid	*C*-terminal side of Trp and Tyr
	Dilute acid	Asp and Pro
	3-Bromo-3-methyl-2-(2-nitrophenylthio-3H-indole (BNPS-skatole)	Trp

115

116 **Establishment of Optimal Digestion Conditions**

117 Factors that impact the effectiveness and reproducibility of digestion of
118 proteins are common to any chemical or enzymatic reactions: pH, temperature,
119 time, choice and ratio of digest enzyme/reagent to substrate/reactant.

120 **pH—** The digestion mixture pH is generally determined by the enzyme or
121 reagent not the protein being analyzed. For example, a highly acidic
122 environment (e.g., pH 2, formic acid) is necessary when using cyanogen bromide
123 as a cleavage agent; however, a slightly alkaline environment (pH 8) is optimal
124 when using trypsin as a cleavage agent.

125 **Temperature—** The optimal temperature is dependent on the cleavage
126 reagent; for example most enzymes have maximal activity in a range of 25° to
127 37°. The temperature may to some degree define the specificity of the enzyme.
128 In such cases the fine tuning of the temperature can be used to optimize the
129 digestion conditions for certain proteins. Ideally, the digestion temperature will
130 minimize sample related chemical side reactions and protein aggregation while
131 maximizing the susceptibility of the test protein to digestion while maintaining the
132 activity of the cleavage agent.

133 **Time—** It is necessary to ensure the digestion is sufficient for intended
134 use to avoid variable digests. A simple time course study is an efficient means to
135 ensure sufficient digestion. Time of digestion varies from minutes to days and
136 aliquots of a single reaction may be appropriately stabilized for analysis to
137 determine the time required for complete digestion of the protein.

138 **Amount of Cleavage Agent—**Sufficient cleavage agent should be used
139 to attain the desired level of digestion within a practical time period (i.e., 2 to 20

140 hours), while the amount of cleavage agent is minimized to avoid its contribution
141 to the peptide map. For an enzymatic digest the protein-to-protease mass ratio
142 between 20:1 and 200:1 is generally used. If the cleavage agent is unstable then
143 the result may be improved by dividing the addition of cleavage agent over two or
144 more additions. Alternately, the use of enzymes bound to a solid support does
145 allow the use of higher relative amounts of protease while avoiding enzyme
146 autolysis contamination. In any event, the final reaction volume should remain
147 small enough to facilitate the next step in peptide mapping—the separation step.
148 Chemical cleavage reagents are usually used in significant molar excess, and
149 may need to be removed at the end of the digestion.

150 **Other Parameters—** The concentration of the test protein should be
151 empirically determined. The test protein concentration should be compatible with
152 the subsequent separation step. The concentration should be low enough to
153 minimize the potential aggregation of intact and partially digested proteins. The
154 pH, composition and ionic strength of the buffer will vary depending on the
155 cleavage agent, the protein and the subsequent method of separation.

156 Digestion can introduce ambiguities in the peptide map as a result of side
157 reactions, such as nonspecific cleavage, deamidation, disulfide isomerization,
158 oxidation of methionine residues, or formation of pyroglutamic groups created
159 from the deamidation of glutamine at the *N*-terminus of a peptide. These
160 degradation mechanisms should be considered during computational modeling of
161 the peptide mapping procedure. Furthermore, autolysis may introduce
162 extraneous peaks produced by the proteolytic enzyme digesting itself. The
163 intensities of autolysis peptide peaks are dependent on the enzyme to substrate
164 ratio and the modifications and quality of the enzyme used. To avoid autolysis,
165 solutions of proteolytic enzyme should be prepared at a pH which inhibits
166 enzyme activity or prepared just before use. Modified enzymes, which are
167 prevented from autolysis may be used. For example, commercial preparations of
168 trypsin (often called "proteomics grade") are available in which the lysine
169 residues of the enzyme have been methylated or acetylated to reduce the
170 number of autolytic cleavage sites. To identify any digestion artifacts a blank
171 determination is performed using a digestion control with all the reagents except
172 the test protein.

173 **Separation**

174

175 Separation of a peptide mixture is meant to resolve its complexity so that a
176 valid interpretation of the data is meaningful and reproducible. Several methods
177 and variations have been published and many are used on a regular basis.
178 Generally all have a common objective to obtain a highly resolved peptide map.
179 Several parameters should be considered when developing methods for
180 useful and robust peptide maps. The complexity of the peptide map will ultimately
181 dictate the optimal set of conditions, column, and mobile phases. Method
182 optimization experiments will be required to obtain the highest quality
183 reproducible chromatogram. The molecular weight of the protein substrate will
184 also influence the complexity of the map and the optimal separation.

185
186 **Chromatographic Separation**
187 Chromatographic analysis of peptide mixtures derived from enzymatic or
188 chemical digestion of recombinant proteins can be used to verify the primary
189 amino acid sequence, locate post-translational modifications, and disulfide
190 linkages (when comparing reduced versus nonreduced preparations) in the
191 protein sequence. Many techniques (e.g., ion-exchange, hydrophobic, and
192 capillary electrophoresis) have been used to separate peptides. However, RPLC
193 is by far the most common method. This method is typically used with high
194 pressure liquid chromatographic systems. In this section, the widely used RPLC
195 method is described as the most common procedure of chromatographic
196 separation of peptide mixtures.
197 **Chromatographic Column-** The selection of a chromatographic column
198 is empirically determined for each protein. Columns with different pore sizes (80–
199 1000 Å) or non-porous based on silica, polymeric, or hybrid supports have been
200 demonstrated to give adequate separation. Column particle size can impact
201 peptide separation. For example, columns with particles sizes <2 µm are
202 available and are typically more efficient than 3–5 µm particles. Generally, octyl
203 or octadecylsilyl bonded phases are ideal for peptides. Octadecylsilane (C18)
204 with 300 Å or smaller pores is the most commonly employed bonded phase.
205 **Mobile Phase-** The most commonly used mobile phase for the LC
206 separation of peptides is water with acetonitrile as the organic modifier; however,
207 other organic modifiers such as methanol, isopropyl alcohol, or *n*-propyl alcohol
208 can be employed. Solvents such as the propyl alcohols may be useful for
209 samples that contain many highly hydrophobic peptides while realizing that
210 hydrophilic or small peptides may elute in a column void peak.
211 Mobile phase additives, such as acids/bases, buffer salts, and ion-pairing
212 reagents, are generally needed to produce high quality chromatographic
213 separations of peptides. The most common mobile phase additive has been
214 trifluoroacetic acid (TFA) where typical concentrations of 0.05% to 0.2% have
215 been employed. Volatile acids and salts can be used to improve MS compatibility.
216 Other additives such as phosphate are less common, primarily due to MS
217 becoming a common detection method, but are useful when separating peptides
218 by LC with UV detection. While TFA has a significant positive impact on the
219 quality of peptide separation, MS sensitivity can suffer with TFA due to ion-
220 suppression. Formic acid, acetic acid, or combinations of these with TFA
221 increases MS sensitivity by reducing ion-suppression. In addition, there are
222 several buffers, such as those containing volatile ammonia, that are compatible
223 with UV and MS detection. Basically, these buffers allow the selection of a broad
224 range of pH values to optimize separation without causing column degradation or
225 compromising UV or MS detection.
226 **Gradient Selection-** Mobile phase gradients can be linear, nonlinear, or
227 include step functions. The separation of highly complex peptide mixtures may
228 benefit from shallow gradients. Regardless, gradients are optimized to provide
229 clear resolution of one or two peaks that will become “marker” peaks to monitor
230 the quality of the test. Isocratic methods (a single mobile phase) are generally

231 avoided for peptide mapping since they provide limited resolution of the peptide
232 mixture.
233 **Other Parameters-** Temperature control of the column is necessary to
234 achieve good reproducibility. Also, the temperature may be used to optimize
235 peptide separation or improve the retention/elution of certain peptides as the
236 resolution typically increases with temperature for a reversed phase column. The
237 flow rate for the mobile phases is based on the column diameter employed. The
238 column chemistry such as pore-size and design and particle size will also
239 influence the optimal flow rate.
240
241 **Detection**
242
243 Several detection methods are available including photometric, such as
244 ultraviolet (UV) absorbance and evaporative light scattering, as well as MS. The
245 choice of detection method will depend on analytical requirements of the peptide
246 map.
247 Ultraviolet absorbance is simple and useful. Methods incorporating LC/UV
248 can be validated and are highly reproducible. The limitation of this method is that
249 UV absorbance provides no peptide structural information. This is a particularly
250 complex situation when peptides co-elute or when unknown peaks appear.
251 Mass spectrometry has become a useful detection method since the LC
252 effluent can be directly introduced into the MS system. With MS compatible
253 mobile phases, direct mass information will correlate to specific eluting or co-
254 eluting peptides. Employing MS allows assignment of masses to the "peaks"
255 eluting from the LC column.
256
257 **Data comparison**
258
259 To determine whether a biotechnology-derived test article is the desired
260 protein product, a peptide map of the test article is compared to that of a
261 reference standard of the desired protein product (See Figure 2). In simple terms,
262 the test article can be sufficiently identified as being the same protein as the
263 reference standard if both sets of peptide map "fingerprints" match. Measures to
264 compare UV absorbance traces may include the number of peaks detected, their
265 relative retention times, the peak responses and the overall elution profile. Visual
266 comparison may suffice, although software approaches to align peaks and
267 provide quantitative comparisons are available. Coupling of the separation to a
268 mass spectrometer for routine analyses provides additional information on
269 comparability between the test and reference samples. This mass information
270 can be correlated to the masses of predicted peptides from the protein digest,
271 giving insight into peptide modifications, truncations, missed cleavages,
272 impurities, and unresolved co-eluting entities under a single peak.
273
274 **ANALYSIS AND IDENTIFICATION OF PEPTIDES**
275

276 *This section gives guidance on the use of peptide mapping during*
277 *development in support of regulatory applications.*
278 Peptide mapping requires the analysis and identification of peptides
279 generated by the enzymatic or chemical cleavage of an intact protein. The aim is
280 to provide a validated method that can characterize at least 95% of the primary
281 structure of the protein. Analysis of peptides can be done directly on-line during
282 chromatographic separation of the peptides, by direct coupling of the
283 chromatography system to a suitable mass spectrometer, or fractions can be
284 collected off-line for further characterization of the peptides within each aliquot.
285 For off-line methods, UV absorbance can be used to monitor the presence
286 of peaks eluting during a high performance liquid chromatography (HPLC)
287 separation. The peptides within each aliquot can be identified using amino acid
288 analysis, Edman sequencing, or MS. When using these identification methods,
289 the sample size may need to be scaled up to provide sufficient material for
290 identification. The chromatography system should be optimized to allow this
291 scaling up so that resolution of the peptides is not compromised. Edman
292 sequencing can be hindered by blockage of the N-terminus of the protein. To
293 overcome this, the blocked peptide can be analyzed using MS or carboxy-
294 terminal sequencing (however, C-terminal sequencing is not commonly available)
295 or using pyroglutamate amino peptidase to remove an amino terminal
296 pyroglutamate if that is the residue blocking the amino terminus.
297 To identify peptides within the peptide map, the raw MS data is analyzed
298 using software that compares the experimental mass-to-charge ratios (*m/z*)
299 obtained using MS with predicted *m/z* values of peptides and MS/MS
300 fragmentation patterns expected from the protein of interest. Furthermore MS/MS
301 analysis can provide the sequence information.
302 For unknown peptides, the experimental data can be compared with data
303 bases using search engines. An enzyme constraint, representing the
304 proteolytic enzyme used to generate peptides from the protein of interest, should
305 be included in the data analysis. An alternative strategy is to compare the
306 experimental tandem MS spectra using data bases and search engines for
307 comparing these types of data. As with comparing experimental and predicted
308 *m/z* values of the intact peptides, dynamic modifications corresponding to the
309 masses of possible post-translational modifications can be incorporated in the
310 data analysis. Depending on the search engine used, scoring thresholds should
311 be established that indicate the confidence in the identification of each measured
312 peptide. An enzyme constraint should also be indicated within the database
313 analysis corresponding to the known proteolytic rules of the enzyme used to
314 cleave the protein of interest into peptides.
315
316
317 **Use of the Peptide Map as an Identity Test**
318
319 While peptide mapping is often viewed as a general methodology, it often
320 involves developing specific maps for each unique protein. Conversely, when
321 similar peptide mapping procedures are applied to different proteins, the set of

322 peptides produced will be distinctly different. A peptide map may be viewed as a
323 fingerprint of that protein. Therefore, peptide mapping can be used to identify a
324 protein.

325 For use as an identity test, peptide mapping relies on the fact that the set
326 of peptides generated by the selective cleavage is dictated by the amino acid
327 sequence of the target protein and the specificity of the reagent used for
328 cleavage.

329
VALIDATION
330

331 For peptide mapping, the intended analytical applications include the
332 evaluation of amino acid sequence, post-translational modifications, and product-
333 related substances and product-related impurities. When utilized for the purpose
334 of identity testing and the evaluation of amino acid sequence and post-
335 translational modifications, the procedure is considered to be a identification test
336 where data are required to support the specificity of the test; however, due to the
337 complexity of the procedure, additional validation data characteristics of precision
338 and robustness must be considered in order to develop meaningful system
339 suitability criteria. Peak identification and evaluation of peak purity (*or* co-elution
340 of peptides) should also be considered.
341

342
Specificity
343 The specificity of the peptide map is established by the comparison of the
344 peptide maps of a suitable reference substance and samples of structurally
345 related proteins. In order to minimize the inherent variability of the test, the
346 procedure is executed on reference substance and samples at the same time
347 and a 1:1 (v/v) co-mixture of the digests from the pre-change and post-change
348 test articles and product reference standard may also be analyzed in parallel
349 using chromatographic separation. Structurally related proteins include those
350 produced by chemical modification of the reference substance and those where
351 the primary sequence differs due to natural species variation or site specific
352 mutagenesis. Chemically modified forms of the reference substance can be
353 produced by exposure to conditions of pH, temperature, or chemical agents
354 known to cause alteration of the primary structure. These alterations typically
355 include deamidation of asparagine and glutamine residues, oxidation of
356 methionine, histidine, or tryptophan residues, and acid catalyzed cleavage of
357 peptide bonds. Peptide maps of structurally related proteins and the reference
358 substance are compared based on pre-determined acceptance criteria. If the
359 peptide map was designed properly, single amino acid substitutions and amino
360 acid side chain modifications should be easily detected by the method.
361 Visual comparison of the retention times, the peak responses (the peak
362 area or the peak height), the number of peaks, and the overall elution pattern is
363 completed initially. It is then complemented and supported by mathematical
364 analysis of the peak response ratios and by the chromatographic profile of a 1:1
365 (v/v) mixture of sample and reference substance digests. If all peaks in the

366 sample digest and in the reference material digest have the same retention times
367 and peak response ratios, then the identity of the sample under test is confirmed.
368 If peaks that initially eluted with significantly different relative retention
369 times are then observed as single peaks in the 1:1 mixture, the initial difference
370 would be an indication of system variability. However, if separate peaks are
371 observed in the 1:1 mixture, this would be evidence of the nonequivalence of the
372 peptides in each peak. If a peak in the 1:1 mixture is significantly broader than
373 the corresponding peak in the sample and reference substance digest, it may
374 indicate the presence of different peptides. The use of computer-aided pattern
375 recognition software for the analysis of peptide mapping data has been proposed
376 and applied, but issues related to the validation of the computer software
377 currently preclude its use in a compendial test. Other automated approaches
378 have been used that employ mathematical formulas, models, and pattern
379 recognition. Such approaches, for example the automated identification of
380 compounds by IR spectroscopy and the application of diode-array UV spectral
381 analysis for identification of peptides, have been proposed. These methods have
382 limitations due to inadequate resolutions, co-elution of fragments, or absolute
383 peak response differences between reference substance and sample fragments.
384 The numerical comparison of the retention times and peak areas or peak
385 heights can be done for a selected group of relevant peaks that have been
386 correctly identified in the peptide maps. Peak areas can be calculated using one
387 peak showing relatively small variation as an internal reference, keeping in mind
388 that peak area integration is sensitive to baseline variation and is likely to
389 introduce error into the analysis. Alternatively, the percentage of each peptide
390 peak height relative to the sum of all peak heights can be calculated for the
391 sample under test. The percentage is then compared to that of the corresponding
392 peak of the reference substance. The possibility of autohydrolysis of trypsin is
393 monitored by producing a blank peptide map, that is, the peptide map obtained
394 when a blank solution is treated with the same enzyme used in the hydrolysis of
395 the protein.
396
397 **Precision**
398
399 To facilitate the determination of the precision of the peptide mapping
400 procedure, an empirical method of quantifying peak responses (peak areas or
401 peak heights) and peak retention factor should be developed. One approach is
402 to make peak response and peak retention time comparisons that are expressed
403 relative to a highly reproducible reference peak within the same chromatogram.
404 The characteristics of the internal reference peak are described below (see
405 *System Suitability, Protein Digestion*). The relative peak response is then
406 expressed as the ratio of the peak response to that of the internal reference peak.
407 The retention factor can be expressed as the retention time of the peak relative
408 to that of the void volume peak. The use of relative comparison methods
409 eliminates the need to make separate corrections for minor variations in digestion
410 and chromatographic parameters.
411

412 **Robustness**
413
414 For peptide mapping, the test parameters include those that affect both
415 the protein fragmentation (i.e., digestion) and peptide separation portions of the
416 procedure. The test parameters associated with these two critical steps in the
417 analytical procedure are considered independently.
418 Factors such as composition of the mobile phase, protease quality or
419 chemical reagent purity, column variation and age, and digest stability are likely
420 to affect the overall performance of the test and its reproducibility. Tolerances for
421 each of the key parameters are evaluated and baseline limits established in case
422 the test is used for routine lot release purposes.
423
424 **Protein Digestion**
425 Many protein fragmentation strategies employ the use of proteolytic
426 enzymes. As a result, the digestion portion of the peptide mapping procedure is
427 inherently more sensitive to minor variation of test parameters. These
428 parameters include digestion pH, buffer, buffer concentration, ionic strength,
429 digestion temperature, digestion kinetics, test article concentration, protease
430 quantity, protease quality, and the stability of the digest. Using a design of
431 experiments approach, these parameters are systematically studied to
432 understand their impact on assay variability. Those digestion parameters where
433 small variations have been shown to impact the precision of the peptide mapping
434 procedure should be carefully controlled within the test procedure using
435 operating ranges established and validated by these studies.
436 **Protease Quality or Chemical Reagent Purity—** A sample of the
437 Reference Standard or Reference substance for the protein under test is
438 prepared and digested with different lots of cleavage agent. The chromatograms
439 for each digest are compared in terms of peak areas, shape, and number. The
440 same procedure can be applied to other critical chemicals or pretreatment
441 procedures used during sample preparation, such as reducing and S-
442 carboxymethylation reagents.
443 **Digest Stability—** The length of time a digest can be kept before it is
444 resolved by the chromatographic method, as well as the conditions under which
445 the digest is stored before chromatography, is assessed. Several aliquots from a
446 single digest are stored at different storage conditions and resolved by the
447 chromatographic method. These maps are then evaluated for significant
448 differences.
449
450 **Peptide Separation**
451 The peptide separation portion of the peptide mapping procedure
452 generally employs standard chromatographic modalities to affect peptide
453 separation. Most notable in this regard is reversed phase HPLC. This section
454 will focus on a discussion of those chromatographic parameters that tend to be
455 more problematic for the peptide mapping procedure.
456 **Column Considerations -** Column-to-column variability, even within a
457 single lot, can affect the performance of the peptide mapping procedure. A

458 reference substance of the protein under test is digested and the digest is
459 subjected to separation using different batches of columns from a single
460 manufacturer. The maps are then evaluated in terms of the overall elution profile,
461 retention times, and resolution according to pre-determined acceptance criteria.
462 To evaluate the overall lifetime of the column in terms of robustness, a
463 peptide mapping test is performed on different columns and the number of
464 injections is varied significantly (e.g., 10 to 250). The resulting maps can then be
465 compared for significant differences in peak broadening and overall resolution.
466 As a column ages, an increase in back pressure might be observed that can
467 affect the peptide map.
468 A sensible precaution in the use of peptide mapping columns is to select
469 alternative columns in case the original columns become unavailable or are
470 discontinued. Analysts should perform a peptide mapping test using equivalent
471 columns from different manufacturers, and examine the maps. Differences in
472 particle shape and size, pore size and bed volume, carbon load, and end-
473 capping can lead to significant differences in retention times and resolution.
474 Modifications in the gradient profile may be required to achieve equivalency of
475 mapping when using columns from different manufacturers.

476 **Temperature=** The chromatographic resolution of peptide maps has been
477 shown to be sensitive to small variations in ambient temperature and that higher
478 temperatures > 45 ^{o} will improve the resolution. The temperature dependence of
479 the separation should be systematically studied to establish an allowable
480 temperature range and the column should be kept within a temperature
481 controlled environment and preferably at a temperature above 45°.
482
483 **Mobile Phase-** The composition of the mobile phase typically used for
484 peptide separations can accelerate the decomposition of the chromatographic
485 matrix. The acidic nature of the aqueous mobile phase (e.g., 0.1% trifluoroacetic
486 acid) has been shown in silica-based columns to result in hydrolytic removal of
487 the end-capping resulting in a loss of resolution. The effects of this phenomenon
488 are most evident from the data derived from column lifetime studies.
489

490 **System Suitability**

491
492 The equivalency between instrumentation used for the validation of the
493 peptide map and for routine quality control testing should be considered. System
494 suitability tests are used to verify that the resolution and reproducibility of the
495 chromatographic system are adequate for the analysis to be done. These tests
496 are based on the concept that the equipment, electronics, analytical operations,
497 and samples to be analyzed constitute an integral system that can be evaluated
498 as such. The application of this definition to the peptide mapping procedure,
499 demands that additional tests, beyond those normally applied to chromatographic
500 separations, be developed to assure that the protein digestion portion of the test
501 has proceeded adequately.
502

503 **Protein Digestion**

504 One approach to developing an appropriate system suitability test to ensure
505 that the protein digestion of the peptide mapping procedure has performed
506 adequately is to identify a highly reproducible peak within the reference
507 substance peptide map. The following criteria should be considered in the
508 selection of a highly reproducible peak.

509 • The peptide is a digestion product shown to be robust with respect to
510 digestion parameters
511 • The peak is chromatographically well resolved
512 • The peak exhibits a good signal to noise ratio
513 • The peptide does not contain any labile amino acid side chain residues
514 • The peptide does not contain sites of post translational modification
515 • The peak signal is free from any other experimental artifacts.

516 Upon identification, an appropriate operating range for the recovery of the
517 highly reproducible peak in the reference peptide map can be developed to
518 ensure that the protein digestion has performed adequately.
519
520 **Peptide Separation**
521 Requirements for the precision of replicate injections of the standard
522 preparation, appropriate peak resolution, and peak tailing should be included.
523 Given the complexity of the peptide map separation, allowable changes to
524 chromatographic operating conditions to meet system suitability requirements
525 must be verified prior to implementation. It is recommended that a reference
526 substance digest be interspersed periodically with test samples to evaluate
527 chromatographic drift during the analysis.
528
529 **Validation Study**
530
531 Validation of peptide mapping requires that a protocol be designed,
532 outlining in detail the experiments to be conducted and the pre-determined
533 acceptance criteria for the peptide map. The validation criteria include
534 robustness, limit of detection, specificity, linearity, range, accuracy, precision,
535 recovery and reagent stability. The acceptance criteria are dependent on the
536 identification of critical test parameters that affect data interpretation and
537 acceptance.
538 The setting of acceptance criteria is based on empirical observations with
539 respect to quantification (peak area or height) and identification (retention
540 factors) for the relevant peaks that comprise the majority of primary sequence of
541 the protein. Examples of appropriate acceptance criteria would be to set
542 resolution between two adjacent peaks that have resolution > 2. For a peak that
543 shows recovery instability set a minimum relative peak area compared to a peak
544 with stable recovery. There are no new peaks in the reference plus sample
545 mixture.

Figure 1. Peptide Map Method Development and Characterization

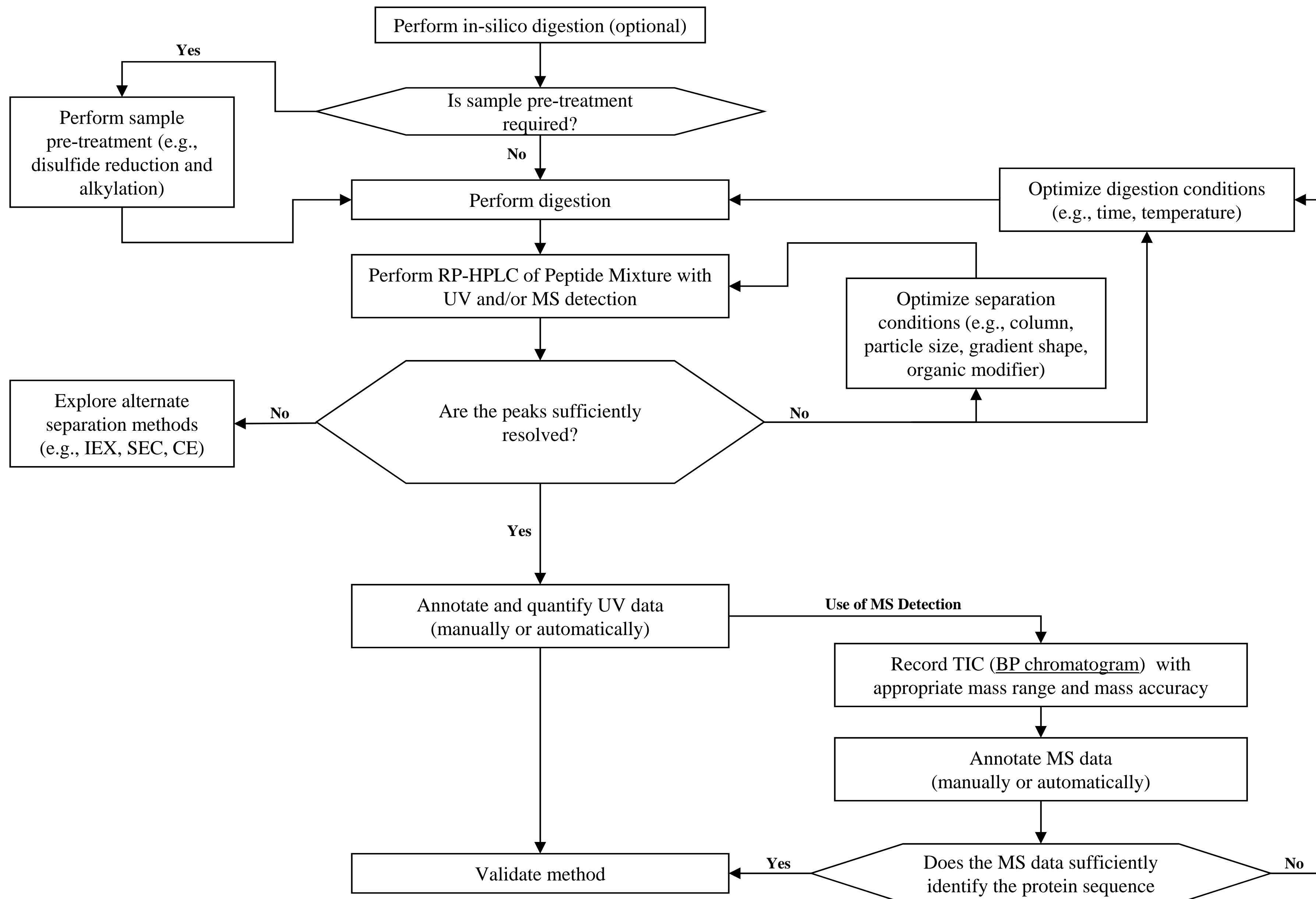

Figure 2. Use of the peptide mapping for identity testing