A Brief Overview of Polypeptide Purification Techniques

Purification of Polypeptides

Polypeptides are a challenging isolation problem[1] due to their wide variety of amino acid sequences, post translational modifications (acetylation, methylation, phosphorylation, glycosylation, etc.) or conformational differences (when the separation mechanism employed does not use denaturing conditions). One needs to consider the source of the material (synthetic or biological), and any limitations due to subsequent steps in analysis or production. The Nest Group's web site, www.nestgrp.com contains an abundant set of applications and instructions for purification which includes reversed phase (RPC), ion exchange (IEX), hydrophilic interaction[2,3] (HILIC), electrostatic hydrophilic interaction[4] (ERLIC) also known as ion-pair normal phase[5], hydrophobic interaction (HIC), and size separation (SEC).

Analytical separations benefit from approaches which reduce the complexity of purity or character of an isolate. Precipitation, solvent extraction, affinity isolation all increase the titer of the polypeptide of interest and reduce the chromatographic burden imposed on the column chemistry of choice. The fundamental mode of separation amplifies the differences between molecules, so that a "one approach fits all problems" constrains one to using "theoretical plates" (N) of a column rather than the more powerful "chemical selectivity" (α, alpha) to obtain or measure a pure component. The third variable in the resolution equation, "capacity factor" or capacity ratio[6] (k') is a measure of the retention of a peak that is independent of column geometry or mobile phase flow rate. It diminishes in effectiveness as it approaches ten column chromatographic void volumes (isocratic), and while not a valid statement, it reflects the change in the slope of a gradient separation when described as the retention time of a component.

Consequently chromatography of complex samples is best done with orthogonal chemical approaches which use distinctly different mechanisms. Use of two of the same fundamental approaches (i.e. RPC-C4 : RPC-C18; Low pH RPC : high pH RPC; HIC : RPC) is less effective at amplifying the differences between mixtures of analytes, than when using truly orthogonal techniques (i.e. IEX : RPC; HILIC : RPC; NP : RPC). The first dimension simplifies the mixture so the second (or third) dimension has fewer competing analytes to separate.Thus a penta peptide mixture of of positional isomers of 4 prolines and one serine may be difficult by just RPC, but HILIC would detect the polarity differences of the position of the serine in the molecule. Similarly, deamidation products result in the positional isomers, aspartic acid and iso-aspartic acid, which are hydrophobicly similar, but differ significantly in polarity. Thus while an RPC isolation of these is difficult, an RPC isolated mixture can be resolved into individual components by ERLIC[7].

One can obtain the maximum resolution orthogonally compared to trying to do it all on a single column, regardless of how efficient the media may be (i.e. sub2 micron UHPLC, or superficially porous, SPP). Components which elute close to one another have similar mechanistic character, but may differ significantly by charge or other functional groups. Use of an orthogonal chemistry can take advantage of this. Whether it is a same charge different sequence elution, or same hydrophobicity but with different polarity, doesn't matter - both are limited by the single chemical approach. A second chemical separation approach will result in a higher probability of a successful analysis or isolation (preparative scale).

Gradient elution is preferred for polypeptide separations

Because large polypeptides diffuse slowly, RPC results in broader peaks than obtained with small molecules. Gradient elution is the preferred approach to minimize analysis times. Gradients are simply condensed isocratic runs. If one were to pump an indefinite amount of solvent, components would elute. So in applications such as trapping or solid phase extraction, it is best to limit the number of column void volumes to prevent unanticipated elution of weakly retained species. The ideal loading/trapping volume is one bed volume less than the k' (elution volume less the void volume) of the first peak of interest. If doing quantitative proteomics, with the goal of detecting and quantifying all polypeptides, it is important to respect this volumetric limit. The safest volume is one chromatographic bed volume of load, but due to detection limits for low copy number polypeptides, this is a major compromise. The solution would be to choose chromatographic conditions which increase the retention of components, or change the column chemistry to achieve the same resulting longer retention.

Bonded phase silicas have a density of roughly 50% so it is easy to estimate chromatographic void volume. It then is the microliter equivalent of the milligrams of packing. Since packing volume is roughly twice that of the mass (in microliters) one can back calculate from the tube volume both the mass and chromatographic void volume (commonly called a "column volume" and not to be mistaken for a tube volume).

Peak widths of polypeptides eluted isocratically are a function of molecular weight. A tri-peptide would produce a peak representing 600 theoretical plates, regardless how efficient the column would be with a small molecule test probes. Gradient elution of polypeptides, even with shallow gradients, is preferred, since it results in much sharper peaks than isocratic elution. HPLC adsorbents are porous particles with a bonded ligand on the silica surface. The hydrocarbon group forming a hydrophobic phase is usually a linear aliphatic hydrocarbon of eighteen (C18), eight (C8) or four (C4) carbons. The length of the hydrocarbon chain often makes little difference in the effectiveness of protein separations. More often it affects the recovery. Polypeptides must enter a pore in order to partition into the bonded phase and be differentially separated. Polypeptides chromatograph poorly, in part because many are too large to enter pores completely which results in a combination of size separation along with the hydrophobic partitioning as evidenced by band (peak) broadening as they diffuse at different rates over different distances. Most polypeptide RPC separations are performed on columns with particles with pores of about 300Å, although some peptides (< ~2,000 MW) may also be separated on particles with 120Å pores and very large polypeptides (e.g. antibody fragments) utilize 1000Å pore columns. Ion exchange in aqueous mobile phases is a slower diffusion process, so 1000Å and larger pores or superficially porous media are employed to improve resolution through shortened diffusion distances.

Adsorption/ desorption approaches for hydrophobic phases.

Peptides desorb from a hydrophobic surfaces in a very narrow organic modifier range due to the precise number of organic modifier molecules required by a polypeptide's conformational structure[8]. This characteristic accounts for the hydrophobic selectivity of very similar peptides in RP-HPLC separations and very sharp chromatographic peaks. RPC has been the easiest approach to peptide separations using a combination of water and acetonitrile (ACN) containing an ion pairing agent such as trifluoroacetic acid (TFA), hexafluorobutyric acid (HFBA) or volatile or non-volatile buffer salts depending on the application and instrumentation employed.

Adsorption/desorption steps take place while the polypeptide is retained on the column. After desorption, very little interaction takes place between the polypeptide and the reversed-phase surface and subsequent interactions have little affect on the separation. Thus column length results in more loading capability with minor increases in resolution once the peptides start to move. Particle size and pore accessibility influence the separation most for a given surface chemistry. The sensitivity of polypeptide retention to subtle changes in the modifier concentration makes isocratic elution difficult because the organic modifier concentration in the micro-environment around the peptide must be changed very precisely for uniform desorption.

For larger polypeptides where native conformation needs to be retained, the use of hydrophobic interaction (HIC) is the recommended hydrophobic technique. Once again, a hydrophobic, surface, but with around ten times less ligand density than RPC columns, is used. Additionally, particles with less surface area (i.e. large pores in the particles) is required to accommodate these larger polypeptide molecules, since they diffuse even more slowly than smaller peptides and thus a closer surface reduces band broadening from a mixed SEC - HIC mechanism. However, rather than using organic solvent to solubilize and adsorb and desorb the protein, a reverse salt gradient is used. The approach is the chromatographic equivalent of salt precipitation of proteins which relies on the dewatering of a protein to induce aggregation and precipitation from an aqueous solution. This is a slower process than in RPC desorption so one obtains broader chromatographic peaks as the salt is removed and the protein comes back into solution.

Because HIC utilizes salt gradients, it should not be used before an ion exchange separation (IEX) since salt promotes elution in IEX. While both are high capacity isolation techniques employing salt or pH gradients, the better combination would be to use IEX as a first step, followed by HIC (then dialysis, if necessary, to remove salt from subsequent steps).

Selectivity differences in HIC are possible within the Hofmeister series of chaotropic salts[9,10]. There are specific Ion effects for these various salts regarding their ability to hydrate surfaces and to selectively retain (precipitate) a mixture of proteins on a hydrophobic surface. Ammonium sulfate is the least selective, although the most used. It is better to incrementally increase and test each effect from the amount of salt used for retention, since different proteins precipitate at different concentrations of salt. Since the rehydration rate is slow, one can obtain a separation that could be lost if the rate of hydration is the determining factor for separation rather than the initial differential solubility.

Diameter and Length Effects on Sensitivity of Detection

The column diameter and length do not affect peak resolution, but do affect sample loading, solvent usage and can affect detection sensitivity under certain conditions. The flow rate and sample loading is proportional to the ratio of the square of the column diameters. Loading is also proportional to the ratio of the column lengths should one be moving

from one length to another. Thus, while decreasing the diameter lowers the amount of solvent used, it can only increase the detection sensitivity, as long as the injection volume/mass is not decreased proportionately too. It is an error to assume that a smaller internal diameter column will of itself increase sensitivity since it is all proportional. However, very small diameter columns are useful when coupling an HPLC with mass spectrometry (LC-MS) when the efficiency of removal of solvent is dictated by the MS design. Tryptic digests of as little as five nanomoles of protein have been separated and collected using 2.1nnID RP-HPLC columns using UV detection. MS sensitivity is greater from its universal, destructive detection mechanism. LC-MS detection using more volatile solvents, such as when doing HILIC, or ERLIC (ion-pair normal phase), can be up to 100 times more sensitive than LC-MS RPC.

Mobile Phases and Temperature Effects

Polypeptides are generally more soluble in aqueous solvents than at high organic concentrations. Increased temperatures aid diffusion by reducing solvent viscosity and increases solubility in these aqueous organic mixtures so sharper, earlier eluting peaks result. However, bonded phases on silica have an upper temperature limit imposed by their covalent bond strength. Thus 40°C is a common temperature, while 60°C approaches the silica-silane bond strength limit.

Desorption from RPC columns is promoted when aqueous solvents contain an organic modifier and an ion-pair reagent or buffer. Elution occurs by gradually matching the concentration of organic solvent molecules to that necessary to cause desorption. The organic modifier solubilizes and desorbs them from the hydrophobic surface while the ion pair agent or buffer controls both the polarity of the column chemistry and that of the polypeptide. It sets the eluent pH and ion pairs with the zwitterionic polypeptides to keep them in a single ionic and hydrophobic form.

Trifluoroacetic acid is the most common ion-pairing reagent for UV based RPC. It is widely used because it has little UV adsorption at low wavelengths and is volatile for easy removal from collected fractions. TFA is normally used at concentrations of about 0.1% (w/v). However, gradients with at a constant concentration of TFA can have a baseline shift when monitoring at 210–220nm resulting from a change in dielectric constant. As the acetonitrile (ACN) solvent environment goes from aqueous to non-aqueous, this affects π-π electron interactions which, in turn, affect the solvent UV adsorption spectrum (190 to 250nm region). To reduce or eliminate this baseline drift, use of 0.1% TFA in Solvent A and 0.085% TFA in Solvent B balances these forces. It is important to use good quality TFA and to obtain it in small ampules. Poor quality or aged TFA has UV absorbing impurities that cause spurious peaks to appear. It is not recommended for LC-MS as it decreases detectability through increased “matrix effects.” Instead, formic acid, and preferably ammonium formate buffer are recommended to control analyte ionic forms and to control eluent pH.

Heptafluorobutyric acid (HFBA) in reversed phase at up to 0.1% v/v dramatically retains peptides containing basic amino acid residues, better than TFA, formic, and acetic acid. However, signal suppression may be significant in LC-MS so UV detection is preferable for both of these ion pairing agents. Unlike TFA one uses more HFBA in Buffer B (Buffer A: 0.54% with Buffer B: 0.60%) to flatten the baseline for UV detection .

Organic acids promote ionization of basic compounds by protonating the C-terminal residue and Asp and Glu side chains enhancing positive ion MS detection. Volatile buffers and ion pairing reagents are necessary to prevent background matrix effects, signal suppression and MS source contamination. Use of buffers is preferable to free acids since they control the ionization and sharpen peaks, especially in HILIC separations. These include: Ammonium formate or acetate (2 - 10mM optimum; one can see suppression effects if > 20mM on column concentration). Acetic Acid at 0.1 – 1.0% v/v. Formic Acid (FA) at 0.1 – 0.5% v/v.

Acetonitrile (ACN) Acetonitrile (ACN) is the most commonly used organic solvent because it is volatile and easily removed from collected fractions; It has a low viscosity, minimizing column back-pressure; It has little UV adsorption at low wavelengths; It has readily available. Isopropanol is often used for large or very hydrophobic proteins. The major disadvantage of isopropanol is its high viscosity and its UV cutoff of 230nm. Adding 1–3% isopropanol to acetonitrile can increase protein recovery, as would using a shorter alkyl chain length column such as C4. Ethanol is often used for process scale purifications for both financial as well as regulatory reasons.

Mobile phases for HILIC and ERLIC separations start at high, aprotic, water miscible solvent concentrations and desorb polypeptides by either an increase in buffer strength and/or by a decrease in the organic solvent component. Polypeptides elute based on polarity and ionic differences[2,3,4,5] from polar column surfaces which range from bare silica to bonded chemistries such as diol, ion exchange chemistries suitable for high organic mobile phases, and an increasingly diverse offering of zwitterionic phases which take advantage of the coulombic interactions of ionic and neutral polar groups in polypeptide mixtures.

The Effect of pH on Peptide Separations

Peptide separations are often sensitive to the eluent pH because of protonation or deprotonation of acidic or basic side-chains. In general, separation at mildly acidic pH (6.0 or 6.5) with triethylamine acetate (TEAA) or phosphate buffer appears to produce better peak shape and resolution, or at least a greater number of resolved components, than separation at highly acidic pH with TFA, TEAP or phosphate. This observation is significant in view of the fact that highly acid conditions such as 0.1% TFA tend to be the norm for reversed phase peptide chromatography. The results shown here suggest that pH 6.0 or 6.5 separations are more likely to separate difficult to resolve components. However, it should be noted that these results are specific to the particular peptide sample used. The conclusions may or may not apply to other samples. A series of trial separations can be helpful in choosing optimal conditions for a specific peptide purification.

Resolution can be changed by using a different ion-pairing agent or a higher pH. Use of TFA is the best starting point for peptide separations. However, consider the use of buffers such as phosphate or hydrochloric acid at low pH or basic pH. Prepare 10-20mM mobile phase solutions of pH 4.4 phosphate, pH 2.0 phosphate and pH 6.5 phosphate. Compare these to 0.1% TFA, 0.1% HCl and 10mM pH 10.0 buffer (on a polymeric or pH resistant column) to find the optimum reagent and pH conditions for a good peptide separation.

Peptide Fragments from a Protein Digest: Deamidation and Oxidation

Protein deamidation results in the conversion of a neutral asparagine to an acidic aspartic acid and its positional isomer, iso-aspartic acid, thus increasing the polarity of the protein. At neutral pH the protein therefore becomes somewhat more hydrophilic. Consequently, use of ERLIC [8] or HILIC would better discriminate between the three forms of peptides rather than by RPC.

Peptide Fragments from a Protein Digest: Glycopeptides

The LC-MS analysis of a tryptic digest provides information about the structure of a protein. However, glycopeptides would co-elute in RPC with non-glycosylated peptides unless a prior fractionation simplified the mixture. A simple method[5] to enrich glycopeptides employs ion-pair normal phase (HILIC) conditions. By adding TFA to the sample and mobile phase one both protonates carboxyls and ion pairs amines of all peptides in a digest, leaving only the glycoforms as the most polar peptides in the mixture. Subsequent HILIC chromatography by HPLC of SPE results in an enriched, selective subset of just the glycosylated forms which can then be LC-MS characterized with RPC or HILIC conditions, as appropriate, or deglycosylated and analyzed by IEX HILIC LC-MS.

References & Resources

1. A Practical Guide to Peptide and Protein Purification for Microsequencing, 2nd Edition, Paul T. Matsudiara (editor), Academic Press, Inc. (1993).
2. Hydrophilic-interaction chromatography for the separation of peptides, nucleic acids and other polar compounds, Andrew J. Alpert, J. Chromatogr. 499 (1990) 177.
3. Review: Hydrophilic-interaction chromatography, Petrus Hemstrm and Knut Irgum, J. Sep. Sci. 2006, 29, 1784 – 1821
4. Electrostatic Repulsion Hydrophilic Interaction Chromatography for Isocratic Separation of Charged Solutes and Selective Isolation of Phosphopeptides, Andrew J. Alpert, Anal. Chem. 2008, 80, 62-76.
5. Identification and Quantification of Glycoproteins Using Ion-Pairing Normal-phase Liquid Chromatography and Mass Spectrometry, Wen Ding, et al., Mol Cell Proteomics 2009 8: 2170-2185. First Published on June 12, 2009, doi:10.1074/mcp.M900088-MCP200.
6. Chromatographic formulas and principles: http://www.lcresources.com/resources/TSWiz/hs210.htm
7. Enhanced Separation and Characterization of Deamidated Peptides with RP-ERLIC-Based Multidimensional Chromatography Coupled with Tandem Mass Spectrometry, Hao, et al. dx.doi.org/10.1021/pr201048c | J. Proteome Res. 2012, 11, 1804–1811
8. Retention Model for Proteins in Reversed-Phase Liquid Chromatography, X. Geng and F.E. Regnier, J. Chrom. 296, 15-30 (1984)
9. Interactions between macromolecules and ions: the Hofmeister series, Yanjie Zhang and Paul S Cremer, Current Opinion in Chemical Biology, Volume 10, Issue 6, December 2006, Pages 658–663. https://doi.org/10.1016/j.cbpa.2006.09.020 .
10. The Hofmeister effect and the behaviour of water at interfaces, Collins KD, Washabaugh MW, Q Rev Biophys. 1985 Nov;18(4):323-422.