The requirements of liquid chromatography (LC) separation systems to separate proteins are often different from those of small molecules as the inherent properties of proteins can result in loss by adsorption, precipitation at zero charge or broad and asymmetric peaks in liquid chromatography. This article looks at the specific separation problems of proteins; the advantages and disadvantages of gel electrophoresis compared with alternative techniques; the potential of pI separations by pH gradients rather than by isoelectric focusing; and describes a two-dimensional liquid chromatography technique that uses pI separation in the first dimension and reversed-phase separation in the second dimension.

Separation Problems Related to Proteins

Samples from living cells, tissue or body fluids can contain thousands of different proteins, depending on the sample and purification method. With expression levels differing by a factor of $10^6$–$10^8$, obtaining a complete picture of the proteins in a particular sample is probably one of the most complicated separation problems of all. Highly selective sample preparation methods, such as affinity chromatography (AC), are often required to reduce the complexity.

Complexity is not the only problem, however. The solubility of macromolecules is not easy to foresee and control. Large hydrophobic membrane proteins have little solubility in aqueous solutions, while the smaller more polar proteins dissolve easily. At the pH of the isoelectric point (pI), where the protein has no charge, the solubility is usually decreased and this can lead to precipitation of the more abundant proteins.
Another feature of macromolecules is the large inherent adsorption energies, resulting in frequent losses by adsorption on solid surfaces as well as on separation matrices such as column materials. Adsorption does not necessarily mean permanent loss but because of the slow kinetics of macromolecules, this may easily lead to overlap between fractions in separation processes. Thus, trapping and separation of proteins by adsorption interactions should always involve as weak interactions as possible in order to counteract the problem caused by slow kinetics.

The peak width of a protein peak in a chromatographic system is also much wider than the peak width of a small molecule, and this is at least partially related to the size of the C-term in the van Deemter equation. This means that if a protein is visualized by a sharp peak in liquid chromatography (LC), this is almost always guaranteed to be a result of gradient elution.

Last, but not least, the peak shape of a protein in a separation system is also dependent on the degree of unfolding of the three dimensional structure. An intact globular protein will be expected to have both a different peak shape and different retention compared with a more or less unfolded structure or a completely denatured structure. There is also a risk of oxidizing thiol groups to disulphides in proteins containing cysteine or methionine groups to sulphotides.

Thus, in our opinion there is no area of separation science with more challenges than in the field of proteins, reminding the reader of the statement of Anderson and Anderson in 1998: “Considered objectively, there is every reason to expect that proteomics will ultimately exceed genomics in total effort, though this effort will sorely be limited by the availability of scientists able to deal with protein’s nonideal properties”.1

Determining the Primary Structure of Proteins
Determining the primary structure (the amino acid sequence) of a protein is now mainly performed by mass spectrometry (MS) or combinations of LC and mass spectrometry. Reductions, alkylations and other derivatizations are occasionally needed as well. After the Matrix Assisted Laser Desorption Ionization Time-of-Flight (MALDI-TOF) mass spectrometers became available, there often appears to be a belief that most structure determinations can be performed with these, which is a misunderstanding. The MALDI-TOF instruments have become very valuable in proteomics because of the open access format, but they cannot be used for sequencing. For this, another mass spectrometer, either a Quadrupole Time-of-Flight (Q-TOF), a triple quadrupole, an ion trap or a MALDI-TOF-TOF is required.2 This also means that a separation system based on the requirement for extraction and transfer of stained spots from 2D GE sheets is inherently more complicated than direct transfer from an LC column to the ion source or to a tray of micro-vials (for MALDI-TOF).

Although 2D GE is the standard reference method in proteomics for separation of proteins, there are many problems with the technique.

Two-dimensional Gel Electrophoresis (2D GE)
In 2D GE the first separation includes isoelectric focusing (IEF) using separation of intact proteins. Today this is performed on an immobilized pH gradient (IPG) strip containing ampholytes to create pH gradients of wide or narrow ranges. After separation into zones by IEF, and treatment with a thiol reducing and sodium dodecyl sulphate (SDS) buffer, the strip is joined with the SDS slab gel. In a buffer containing high concentrations of chaotropic agents such as urea for unfolding the proteins, thiourea and other additives to prevent oxidation of disulphide groups and reduce disulphides and detergents to keep the proteins in solution, electrophoresis is then performed on polyacrylamide gel, separating the proteins. SDS complexes according to size only.

Some proteins can be lost by adsorption to the IPG matrix, the presence of 2M thiourea and 5–7 M urea has been recommended in the first dimension too.3

Although 2D GE is the standard reference method in proteomics for separation of proteins, there are many problems with the technique. Inherent problems are the lack of good quantification and of high reproducibility.4

There is also the solubility problem in the first dimension, particularly with hydrophobic proteins. Partial precipitation may lead to “smearing” of bands along the flow direction. Aggregation of proteins to larger clusters adds to that problem and lowers the resolution even more. Possibly the most serious problem is the lack of detection of the minor components because the staining procedures result in a limited dynamic range of detection. Thus, the number of proteins actually

![Figure 1: Capillary pH-gradient ion exchange chromatography of α-lactoglobulins B (1) and A (2) using different injection volumes. The pH-gradient and the recording of the chromatograms were started after the sample was loaded. Column: 10 cm × 0.32 mm i.d. PL-SAX (10 μm) column. Start buffer (A): 10 mM piperazine and 10 mM N-methylpiperazine, pH 6.8, eluting buffer (B): same as start buffer, but pH 4.3. Gradient programme: 0–100% B in 25 min, then 100% B for 5 min. Flow-rate: 8 μL/min. Sample concentration: 5, 0.5, 0.05 and 0.005 mg/mL of each protein (diluted in start buffer). Column temperature: 30 °C. UV detection at 280 nm. From ref. 13, with permission.](www.lcgceurope.com)
identified by 2D gels is very low. For three bacterial species, less than 5% of the proteins in SWISS-PROT were identified on the gels. Within the 5%, only between 5 and 15% were hydrophobic, compared to the theoretical hydrophobic contents of 16–29%.6

Also, biomarkers for diseases, in serum, are often present at concentrations below 10 ng/mL, which makes the conventional methods inadequate for detection.6 In a human cell the most abundant protein is often actin, with a concentration of 10^8 molecules per cell, while some cellular receptors or transcription factors are present at only 10^2–10^3 molecules per cell.7

**Alternatives to 2D GE**

Depending on the type of sample and the amount of protein available, there are at least two main alternative routes. With plenty of sample at hand, off-line separations are simpler and often an advantage as the first dimension. Isoelectric focusing can separate proteins according to pI values and has many advantages as a prefractionation technique.7 However, with very limited sample amounts, on-line techniques allow transfer of the whole fraction to the second dimension, and with miniaturized columns, little band broadening is obtained. Unfortunately, isoelectric focusing is not easily included in an online two dimensional instrumentation with microbore columns.

In samples containing relatively few proteins, the “shot gun” approach is an alternative, by enzymatic degradation of the mixture of proteins to peptides and coupled LC–MS–MS or MALDI-TOF MS analyses of the peptide mixtures for protein identification. Even with some overlap of peptides, the protein databases have the potential of identifying the majority of the proteins, depending on the concentration.

Samples containing more complex sources of proteins always benefit from a pre-separation of the proteins prior to enzymatic degradation, either in one dimension or in two dimensions.

With a 2D system, the two separation principles should be orthogonal, such as a combination of pI separation and reversed-phase chromatography or hydrophobic interaction chromatography. For peptides, a combination of cation exchange and reversed phase is common today.

The separation of proteins on cation exchangers with salt gradients has some disadvantages; foremost that there is no direct relationship between the pI and the salt concentration and the retention time. Eluting proteins by a pH gradient would be much more advantageous, since the pH of each eluting peak in principle can be measured directly, resulting in a direct link between eluting pH and pI of the proteins. Unfortunately, common pH gradients always have been known for low repeatability and non-linearity.

Thus, an important question is whether a separation according to pI can be implemented as the first dimension of a 2D LC separation.

**Separations According to pI**

In the late 1970s, Sluyterman and co-workers8–9 described a pH-gradient ion exchange chromatography (IEC) method that used the buffering capacity of the column to generate linear pH-gradients without external mixing of the buffers, which they termed chromatofocusing (CF). CF is most commonly performed by titration of a weak anion-exchange column, (e.g., DEAE(diethylaminoethyl)-functionalyzed cellulose), which has initially been equilibrated with a start buffer at a pH higher than the pI of the most basic proteins, and is eluted with a buffer of low pH. To obtain as linear pH-gradients as possible, and larger peak capacity, the buffers usually contain mixtures of polyampholytes with different pKa values to get an even buffering capacity over the chosen pH range. However, CF has several limitations; limited reproducibility because of batch-to-batch variability in the chemical composition of polyampholyte mixtures, formation of

---

**Figure 2:** Generation of pH-gradients on a 10 cm × 0.32 mm i.d. PL-SAX (10 µm) column. Start buffer (A): 10 mM piperazine; 10 mM N-methylpiperazine and 10 mM imidazole, pH 8.5, eluting buffer (B): same as start buffer, but pH 4.0. Gradient programme: 0–100% B in (a) 10 min; (b) 20 min and (c) 30 min. Flow-rate: 8 µL/min. Column temperature: 30 °C. UV detection at 280 nm. From ref. 13, with permission.

**Figure 3:** Capillary pH-gradient IEC of diluted skimmed bovine milk (1:10, v/v) on a 10 cm × 0.32 mm i.d. PL-SAX (10 µm) column. Start buffer (A): 5 mM piperazine and 5 mM N-methylpiperazine, pH 6.8, eluting buffer (B): same as start buffer, but pH 4.3. Gradient programme: 0–100% B in 25 min, then 100% B for 5 min. Flow-rate: 6 µL/min. Injection volume: 5 µL. Column temperature: 30 °C. UV detection at 280 nm. The pH-gradient and the recording of the chromatograms were started after the sample was loaded. From ref. 13, with permission.
association complexes with proteins, high background absorption with UV detection and little flexibility in choice of buffer concentrations.

While CF generally exploits the buffering capacity of the ion exchanger to obtain a retained intra-column pH-gradient, this is, however, not necessarily a problem, as long as it is sufficiently robust in the long run. Another smaller electrode was built into a housing for microlitre flow but the response time of the pH electrode was much too long. Thus, we decided to use a larger volume pH electrode in order to give fast response, requiring 1–2 mm i.d. ion exchange columns in the second dimension, on a PS/DVB monolith, with TFA in the mobile phase. They used ion-exchange materials with small buffering capacities in the applied pH range in combination with buffer components that gave an enhanced flexibility in controlling the slope of the pH-gradient. In contrast to CF, pH gradient IEC also allows the use of higher buffer concentrations without affecting the slope of the pH-gradient. Accordingly, improved chromatographic performance can be obtained with only a few common buffer components, while still attaining the characteristic focusing effect of the protein bands.

Recently, Andersen et al. developed a pI separation of proteins on packed capillary columns, using a strong anion exchanger, which improved the peak shape of the proteins while maintaining the focusing effect (Figure 1). In the weakly basic to acidic pH range, a linear pH curve was obtained (Figure 2), which could be used to determine β-lactoglobulins in skimmed milk (Figure 3). The pH monitoring was performed by a flow-through cell with a pH electrode. Unfortunately, the cell was not sufficiently robust in the long run. Another smaller electrode was built into a housing for microlitre flow but the response time of the pH electrode was much too long. Thus, we decided to use a larger volume pH electrode in order to give fast response, requiring 1–2 mm i.d. ion exchange columns in order to avoid too much peak broadening.

It is a basic fact of this kind of pH gradients that a linear pH curve is only obtained in a range outside the buffering area of the ion exchanger. Thus, starting with strongly basic pH led to a non-linear pH curve because of titration of the ion exchanger. This is, however, not necessarily a problem, as long as it is reproducible (Figure 4). A mixture of acidic and basic proteins in a wide range of pI values could easily be separated in two groups by a wide range pH gradient (Figure 5).

In order to avoid precipitation of hydrophobic proteins at their pI values and to reduce potential losses by adsorption, trifluoroethanol was added to the mobile phase in 2–20%, but with disastrous results. The peak shape of the proteins was strongly malformed with strong peak broadening. Replacing trifluoroethanol with ethanol, however, gave more symmetric and narrow peak shapes. Addition of 20% or more of ethanol dehydrated the pH electrode, but 10% could be used without problems, as shown in Figure 5.

2D LC of Proteins
In the second dimension, if compared to 2D GE, the proteins are intended to be separated by size. Because size exclusion chromatography (SEC) is a method with a very narrow separation window, with little ability to separate multiple components, reversed-phase chromatography is variably chosen to represent a separation according to size. Zhu et al. used non-porous 1.5 µm silica based C18 beads in the second dimension of their off-line 2D method for proteins, after chromatofocusing. The absence of oxidation of methionine groups to sulphoxides was reported as an important advantage of the LC method, compared to 2D GE.

In the present article, large pore polystyrene-divinylbenzene (PS-DVB) materials were selected in order to avoid silanol interactions on silica-based columns and to mimic some of the properties of monolithic columns, since even large membrane proteins have been demonstrated to elute without problems on PS-DVB monoliths, with TFA in the mobile phase. Reversed-phase columns do not separate according to hydrophobicity/size exclusively, but within a group of proteins already selected for their acidic/basic properties, a size-related separation is a reasonable expectation. Figure 6 demonstrates the results of connecting a pH gradient in the first dimension with a reversed-phase separation in the second dimension, on a PS/DVB column. The PS/DVB particles had a pore size of 4000 Å to reduce band broadening caused by slow kinetics, to reduce adsorption in narrow pores and to avoid exclusion of large proteins from narrow pores. As a result of the low back pressure on the wide-pore particles, high flow-rates allowed short elution times. The timescale in the figure is the total time of analysis.

The basic fraction ends at 14 min and the acidic fraction starts at 20 min. The separations of the basic proteins in the second dimension were according to MW. The acidic proteins were also separated according to size except that α-lactalbumin (MW 14 200) eluted after the lactoglobulins (MW 18 400) and BSA eluted prior to ovalbumin. This demonstrates that expectations of MW related order of elution on the PS/DVB column must be treated with care.

Conclusions
Separation and focusing of proteins can be obtained by pH-gradient ion-exchange chromatography when the pH of...
Figure 5: Acidic and basic proteins separated in groups by pH-gradient ion exchange chromatography on a 250 × 0.32 mm i.d. PL-SAX (10 µm) column. The acidic proteins (ovalbumin, α-lactalbumin, bovine serum albumin, β-lactoglobulin –A and –B, insulin) eluted at pH 4.0–4.6, and the basic proteins (myoglobin, cytochrome C, α-chymotrypsinogen, lysozyme) eluted at pH 8.3–9.0. Start buffer (A): 10 mM piperazine, N-methylpiperazine, imidazole, trizma, diethanolamine and ethanolamine in water, ethanol (90:10) (pH 11.1), eluting buffer (B): same as start buffer, but pH 2.9. Gradient programme: 0–100% B in 10 min, then 100% B in 20 min. Flow-rate: 5 µL/min. Injection volume: 1 µL. UV detection at 280 nm.

Figure 6: Two-dimensional separation of proteins including separation by pH-gradient ion exchange chromatography on 10 µm PL-SAX (250 × 0.32 mm i.d.), trapping of two fractions on short 5 µm PLRP-4000A columns (5 × 0.32 mm i.d.) and backflush and separation on 5 µm PLRP-S-4000A (100 × 0.32 mm i.d.).
1. Dimension: as in Figure 5.
2. Dimension: Mobile phase A: 0.1% TFA in water. Mobile phase B: 0.1% TFA in acetonitrile. Gradient programme: 10–30% B in 1 min, then 30–50% B in 10 min, then 50–80% B in 3 min. Flow-rate: 10 µL/min. Elution order: 1 = Cytochrome C, 2 = lysozyme, 3 = myoglobin, 4 = α-chymotrypsinogen, 5 = insulin, 6 = β-lactoglobulin A and 7 = β-lactoglobulin B, 8 = α-lactalbumin, 9 = BSA and 10 = ovalbumin.

Acknowledgements
The project has been supported by the EU research training network HPRN-CT-2001-00180 (COM-CHROM).

References

Tyge Greibrokk is professor of analytical chemistry and is currently also Head of the Chemistry Department at the University of Oslo, Oslo, Norway.
Elsa Lundanes is also professor of analytical chemistry in the Chemistry Department.
Milaim Pepaj is a PhD student of analytical chemistry in Oslo.
Katerina Novotna is a PhD student of analytical chemistry at the University of Pardubice, participating in an EU-research training network coordinated by the Oslo group.
Thomas Andersen is a former PhD student of the Greibrokk/Lundanes group and is now a researcher at the G&T Septech company in Oslo, Norway.