Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Chromatography B, 815 (2005) 11–24
JOURNAL OF CHROMATOGRAPHY B
www.elsevier.com/locate/chromb

Review

Peptidomics, current status

Mikhail Soloviev*, Paul Finch

School of Biological Sciences, Royal Holloway, University of London, Egham, Surrey, TW20 0EX, UK

Received 2 September 2004; accepted 1 November 2004
Available online 2 December 2004

Abstract

Characterisation of the complement of expressed proteins from a single genome is a central focus of the evolving field of proteomics. Traditional proteomics technologies were developed in the 20th century and are based on two-dimensional electrophoresis or multidimensional liquid chromatography. These facilitated functional genomics analysis, but they currently represent a significant bottleneck to progress in this area. We are now witnessing the development of novel alternative technologies for use in expression proteomics research. This review aims to familiarise the reader with the principles underlying the peptidomics approaches to proteomics research and provide examples of their applications.

Keywords: Peptidomics; Reviews

Contents

1. Introduction

Proteins were first purified in the 19th century, nearly 100 years earlier than scientists learned about nucleic acids. For the last 50 years, we have lived through the revolution in Molecular Biology research, when the science advanced from having solved DNA structure (J.D. Watson, F.H.C. Crick and M.H.F. Wilkins in 1953, Cambridge, UK [1]), through to DNA sequencing (F. Sanger, MRC LMB, Cambridge, UK [2]), human genome sequencing in 2001 [3,4] and high-throughput approaches for the analysis of individual genomes

* Corresponding author.
E-mail address: mikhail.soloviev@rhul.ac.uk (M. Soloviev).

1570-0232/$ – see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.jchromb.2004.11.011

being developed by Solexa Ltd. (Little Chesterford, UK) (www.solexa.co.uk). If protein research was carried out at a similar pace, we should have been able by now to determine protein composition of every cell of every organism and perhaps in real time mode. But, this is still science fiction. Many researchers are still reliant on two-dimensional electrophoresis, spot identification and gel excision to obtain that one protein of interest, which will help solve that one problem. In the last few years, a variety of novel technologies have sprouted around that cornerstone of protein research—the studying of *proteins*.

2. Proteomics

Proteomics can be defined as the systematic analysis of proteins to determine their identity, quantity, structure and function. Until recently *proteins* were studied using either a single or a combination of established techniques, which could be generally divided into *separation* techniques and *identification* techniques. The former includes *chromatography*, *isoelectric focusing*, *electrophoresis* and their combinations (e.g. two-dimensional electrophoresis). Following their *separation*, proteins can be identified using various approaches either indirectly (e.g. by the size on a gel, through a functional assay, or a ligand binding assay or by affinity reagents) or directly, through protein sequencing (from the N-terminus by Edman degradation or from the C-terminus using carboxypeptidases or thiocyanate degradation, etc.) or using mass spectrometry (whole proteins or proteolytic peptides, e.g. by mass matching or de-novo sequencing). Traditional approaches to *protein* separation rely on the use of a number of consecutive purification stages, for example iso-electric focusing followed by SDS-PAGE electrophoresis, one chromatography separation followed by another (e.g. ion exchange followed by reverse phase chromatography, etc.). Often each separate purification stage would be followed by identification and quantitation stages. Consequently, the majority of traditional proteomics techniques do not allow a highly parallel approach due to their physical limitations, and because of their cost, poor reproducibility and large sample consumption. That is why new truly multiplex approaches for protein research are urgently required.

A few years ago, “protein microarrays” emerged as a potential substitute to one and a half century old protein analysis techniques. Protein microarrays (in many cases the miniaturised arrays of antibodies) are necessarily miniaturised antibody dot-blots and a direct import of the DNA array approach applied to proteins (and chosen for its simplicity, speed and capacity for quantitation). Earlier, Pat Brown and colleagues have shown [5] that elements of DNA chip technology (glass slides, Cy-dyes, pin spotting) can be transferred to create a functioning protein expression chip. Unlike DNA “chips”, protein microarrays exist either as direct binding or sandwich-type (ELISA) assays. Direct binding assays in microaray format revolutionised molecular biology, but often failed when applied to proteins. In a direct binding assay the unknowns, i.e. proteins being assayed, are labelled directly with a detection reagent (i.e. a fluorophore) and assayed by binding to the immobilised antibodies. Unlike DNA microarrays, for which the relative abundance of cDNAs can be measured directly by a fluorescent scanner, without further processing of the chip, protein abundance is not a simple function of the signal [6]. Due to a wide range of antibody affinities to their antigens (in contrast to DNA–DNA interactions, which have an inherently narrower range of “affinities” which are also easy to estimate), quantitation of proteins using arrays of antibodies (or indeed any highly parallel affinity assay) ideally requires the use of competitive assays [7]. The use of sandwich assays negates the necessity for labelling the proteins of interest with a detection reagent and results in higher detection sensitivity, but the heterogeneity of antibody affinities remains a problem.

Compared to earlier days, the modern proteomic tools often rely on highly parallel analysis, miniaturised and chip-based technologies. However, most of the modern developments in the area of proteomics, including chip-based proteomics, are based on traditional and established *protein* purification and separation techniques. So far chromatography remains the most widely used tool in protein analysis (over 15,000 publications in 2003 alone exist in the PubMed National Library of Medicine database, http://www.ncbi.nlm.nih.gov/entrez), followed by various modes of electrophoresis (over 12,000 papers in 2003, same database). Other more recent additions in the range of the proteomics tools include mass spectrometry [8,9] and surface plasmon resonance (SPR) based techniques [10]. Mass spectrometry analysis of proteins is widely used in identification of proteins through mass matching of their proteolytic peptides or de-novo sequencing. It is often carried out using either matrix-assisted laser desorption ionisation-time of flight (MALDI-TOF) technique followed by database mass matching or using MS/MS techniques with collision-induced dissociation to further fragment peptides enabling an amino acid sequence to be generated. SPR-based detection has been exploited commercially by BIAcore AB (www.biacore.com), XanTec bioanalytics GmbH (www.xantech.com) Genoptics (www.genoptics.fr) and Applied Biosystems (home.appliedbiosystems.com). Sensitivity of protein detection approached low-picogram range (for BIAcore, [11]) and is close to that obtained with fluorescent-based affinity assays. SPR is capable of detecting unlabeled sample and is therefore advantageous over ELISA or fluorescent-based assays. BIAcore platform (to date the most widely used of all the SPR machines) combines robotic liquid handling with a miniaturised and automated affinity assay platform, but it is nevertheless, the same “old” affinity assay prone to the same problems. So what are these problems that prevented protein sciences for more than a century from advancing at the same rate as the studies of nucleic acids?

2.1. What is wrong with proteins?

There is little in common between the two most abundant biological polymers—nucleic acids and proteins. Yet, researchers from all over the world are trying to apply nearly identical techniques for their separation and analysis. Proteins and nucleic acids represent two completely different classes of biological polymers and should (but are not) treated as such. For example, mRNA (the main target in "transcriptomics") is usually extracted from the cytoplasm (and more rarely from the nuclei). mRNAs are highly soluble, polar molecules which can be easily separated from the rest of the cell, converted to cDNA molecules with more uniform lengths' distribution and enzymatically amplified if required prior to an "affinity" assay on an oligonucleotide array (usually an array of short DNA fragments of identical lengths and similar chemical composition) or cDNA array. Unlike mRNAs (or even genomic DNAs) proteins are found in different cell compartments (cytoplasm, a range of intracellular organelles) or as secreted extracellular proteins (in various body fluids). Furthermore, proteins range from highly soluble hydrophilic proteins, to membrane associated and transmembrane proteins containing multiple hydrophobic transmembrane domains. Moreover, proteins often exist as multisubunit complexes or can form large macroscopic complexes with other proteins (e.g. a postsynaptic density, where a large number of transmembrane and soluble proteins are associated with the membrane and each other forming complex and functional protein network). No single optimised conditions exist to suit a wide range of protein physical and chemical properties. When nucleic acids are purified from different samples it is assumed that all different mRNAs (or DNAs) are extracted to the same or equivalent degree and such extraction is reproducible between samples. It would be naive to believe that *all* cellular proteins can be solubilised and extracted efficiently and that such extraction can be reproducibly repeated for many different samples. This means that protein composition of two different tissues (for example brain and muscle) cannot be compared quantitatively even if suitable affinity assays were available, since inherent variability at the protein extraction step will make it impossible. The second major difference between nucleic acid and protein affinity assays is in the heterogeneity of the "affinity reagents" used. With nucleic acids (whether used in a blot format or as a microarray), the binding signal is proportional to cDNA concentration (with only narrow range of affinities of complementary strands interaction) and the assay should be suitable for measuring relative or absolute mRNA/cDNA abundances in the sample. Protein affinity reagents (i.e. antibodies or their fragments) have a much wider range of affinities and (unlike nucleic acids where affinity depends on the base composition and can be accurately estimated) these cannot be predicted for antibody–antigen interactions. If affinities of interaction differ for different antibodies used, the signal (antigen binding) will not be a simple function of the antigen concentration as it is generally the case with nucleic acid assays.

Another major difficulty facing proteomic scientists is the availability of affinity reagents. Nucleic acids are both information carriers and affinity reagents. Knowledge of a protein (i.e. antigen) sequence, on the other hand, does not necessarily mean the availability of an affinity reagent and/or information on its affinity and specificity. Currently, there are no complete sets of affinity reagents available with the required specificities and affinities. This is due firstly, to the absence of information on exactly which protein and which post-translational modification pattern is expressed differentially under particular conditions; and secondly, a variable success rate and significant time requirements for development and characterisation of each antibody, unlike the synthesis of an oligonucleotide or purification of a PCR product. Furthermore, not all proteins can be used as antigens directly. For example, multi-transmembrane proteins are neither easy target for extraction nor suitable for immunisations. Often fragments corresponding to hydrophilic parts of antigens have to be used instead of the antigens, this can affect affinity and specificity of antibody–antigen recognition and further impedes their development. It is unlikely that a generic set of affinity reagents against all proteins (even from one species) will be available in the near future.

Yet another major difference is analyte stability. Nucleic acids are assayed in a "denatured" form and degradation of the molecules will be well tolerated in a hybridisation assay. Protein antigens often interact with antibodies through a 3D interface formed by non-adjacent residues, thus both degradation and denaturation may lead to the loss of antibody–antigen recognition.

3. Peptidomics: the way forward

Methods which work well with DNA (e.g. quantitation using oligonucleotide microarrays) may fail if applied to proteins. Mass spectrometry techniques (mostly MALDI-TOF MS and TOF–TOF MS)—the great driving force behind the recent progress in proteomics—speeded up *proteomics* similarly to how DNA sequencing accelerated Molecular Biology. However, mass spectrometry (not DNA sequencing) remains a serial technique (where individual samples are analysed sequentially) and despite all the progress to date provides a significant barrier to surmount. A parallel affinity assay (e.g. in a microarray format) remains, therefore, a favourite option for a truly high-throughput *proteomic* analysis. Main problems associated with such assays are: *protein extraction* from different tissues, cells and sub-cellular organelles (and solubilisation where necessary); *heterogeneity of the proteins'* physical–chemical properties; *availability of affinity reagents* and *availability of antigens* (sufficient amounts and purity) to generate affinity reagents; heterogeneity of *antibody affinities* and problems associated with *protein denaturation and degradation*. How can these hurdles be overcome?

- *Heterogeneity of the proteins'* physical properties is an inherent nature of *proteins* and cannot be dealt with unless proteins are digested (in a fully predictable manner) and short peptides are released. On average, just over 50 tryptic peptides will be obtained from each protein in the human proteome if trypsin digestion is used (based on protein lengths and trypsin digestion frequencies).
- *Protein extraction* from different tissues, cells and subcellular organelles cannot be avoided if *proteins* are sought. However, material can be extracted from samples without protein solubilisation, for example by treating samples with trypsin [12]. Such treatment will destroy proteins, but will release a large number of peptides (mostly hydrophilic, since hydrophobic fragments will remain inside the lipid membranes). Moreover, when a complex protein sample is digested (with trypsin), the distribution of hydrophobicities of individual components, i.e. proteins (prior to digestion) or peptides (following the digestion) changes dramatically. Consider for example an adenosine A1 receptor [13]. This is a 326 amino acid long 7TM receptor, which is a strongly hydrophobic molecule (Kyte and Doolittle hydrophilicity index [14] is −0.68). Out of the 326 amino acids of the A1A precursor protein, 173 amino acids (53.1%) are strongly hydrophobic (these form seven transmembrane helixes and the signal peptide). However, upon tryptic digestion (without prior purification, i.e. of the membrane associated receptor), five tryptic peptides (of six amino acids or longer) are predicted to be produced from the receptor: *VNQALR*, *TVVTQR*, *ANGSVGEPVIK*, *IWNDHFR* and *CQPKPPIDEDLPEEK*. Their respective hydrophilicity indexes are 0.28, 0.17, −0.05, 1.19 and 1.63. Protein and peptide hydrophilicity correlates well with their antigenicity [15]. Four of these A1A peptides are hydrophilic and two are strongly hydrophilic (and therefore very likely, immunogenic). In addition to the better yield and more straightforward extraction procedures, the peptidomics approach should yield a more immunogenic population of peptides. Because, on average, more than 50 peptides are generated from each human protein following a tryptic digestion, it is easier to find a suitable hydrophilic (immunogenic) peptide to generate anti-peptide affinity reagents (see also below).
- *Availability of affinity reagents* and *availability of protein antigens* to generate affinity reagents—is a bottleneck in affinity proteomics. It is estimated that currently an order of 50,000–100,000 various antibodies might be available worldwide representing ca. 5,000–10,000 different genes/proteins. But even this comes short of an estimated 400,000 proteins and isoforms potentially existing in each individual human cell proteome. Moreover, not every protein sequence is highly similar between different species, making antibodies developed against, for example, human targets unsuitable for use with mouse proteins and vice versa. Generating traditional antibodies suitable to assay native proteins is time consuming and requires that sufficient amounts of appropriate antigens are available in their native non-denatured state. Anti-peptide antibodies were made before (in cases where no antigen could be purified or no full-length sequence was available) but were not used widely, mostly because the anti-peptide antibodies often fail to recognise properly folded protein (though may often work on Western blots where antigens are denatured). Such antibodies (whether traditional IgGs or antibody fragments produced by phage display [16], ribosome display [17], mRNA display [18] or other molecular display technologies [19] become very useful if one aims to assay proteolytically digested protein samples. For example if the sample is to be digested with trypsin, one would need to predict tryptic digestion peptides in silico, chose the most immunogenic peptide(s) and order anti-peptide antibodies. A number of suppliers offer affordable polyclonal or monoclonal antibodies e.g. Eurogentec Proteomics (www.eurogentec.co.uk) or Sigma-Genosys (www.genosys.co.uk). The availability of peptides (note, that no antigen purification is required, just the sequence !) and their affordability makes this route very attractive for low volume users and manufacturers (tens to hundreds of targets). Higher throughput analysis (thousands to hundreds of thousands) should consider using phage display or other display technologies [20].
- Heterogeneity of *antibody affinities* is a problem for both anti-protein and anti-peptide antibodies (though less so for the latter). The use of competitive affinity assays (which can compensate for the heterogeneity of antibody affinities [6]) should cure this problem. The use of anti-peptide antibodies here is preferred over traditional anti-protein affinity reagents due to the availability of "antigens" (i.e. synthetic peptides) for competition style of assays.
- *Protein denaturation and degradation* may present a real problem in *protein* affinity assays, where preservation of the antigenic determinant is crucial. However, this ceases to be an issue in anti-peptide affinity assays where antigenic epitopes are formed mostly by short linear polypeptide fragments with little or no secondary structure.

The huge range of physico-chemical properties of the protein makeup of even the simplest organism means that either affinity reagents must be created and optimised for every protein of that proteome in some as yet undiscovered buffer OR the proteins could be made more homogeneous in their physico-chemical properties. Clearly, this second option is more favourable. In peptidomics, each protein is broken down into many smaller components, resulting in the availability of a large range of peptides thus allowing *multiple independent assays* for the same protein target to be performed (similarly to Affymetrix DNA oligonucleotide arrays, where up to 20 oligonucleotides may be generated against the same mRNA sequence http://www.affymetrix.com), thus increasing the reliability of the assay. Peptidomics enables a high-throughput screening of proteins in a microarray format and has several advantages over the affinity capture of intact proteins. As peptides are much more stable and robust than proteins, protein degradation is not an issue since

only one or a few intact peptides would be required for the analysis. Peptides are also particularly suited for detection by mass spectrometric techniques, such as MALDI-TOF MS for direct analysis of samples on a solid substrate such as microarrays. The peptide mass range is such that isotopic resolution is easily achieved and hence their masses can be accurately determined, allowing for mass matching database searches to be performed to confirm the specificity of the affinity capture. Digestion of cellular fractions or even intact tissues results in the release of peptides, which will be mostly hydrophilic, thus further improving the assay. Also, antibody can be against linear unfolded fragments not native folded proteins and therefore "antigens" can be more easily generated, such as by chemical synthesis of in silico predicted peptides against which antibodies are raised. Such affinity reagents can be obtained in a truly high-throughput manner and against most antigenic peptides, and their specificities and affinities can be more easily controlled.

4. Affinity peptidomics

Three years ago, the peptidomics approach was introduced [21], in which the composition of a protein mixture is determined by directly assaying the peptides from crude tryptic or otherwise digested protein preparations, instead of full-length native protein preparations. In the more recent report by Scrivener et al. [22], tryptic peptide sequences derived from a vascular cell adhesion molecule (VCAM) were used to generate recombinant affinity reagents from a phage display library. Tryptic digestion of the VCAM results in nearly 70 fragments, of which nearly 50 are hydrophilic. From these, five peptides were chosen for anti-peptide antibody development (single chain Fv(s), Table 1). Two to three different clones were generated for each of the peptides and were tested for binding respective peptides. Single chain antibodies were immobilised on hydrogel-coated silicon slides, incubated with respective antigens, and detected directly by MALDI MS (Fig. 1). Table 1 summarizes the results reported by Scrivener et al. [22]—the best response was obtained using affinity reagents developed against the most hydrophilic peptides (Kyte and Doolittle hydrophilicity 0.88, 0.96 and 1.44, respectively, Kyte and Doolittle [14]), whilst less hydrophilic peptides resulted in weaker scFvs or no binding at all (0.58 and 0.35, respectively).

The majority of proteins from the human proteome (Swiss-Prot 40.27) have overall predicted hydrophilicity of 0.45 (Fig. 2) and differ in their folding, physical properties and cellular localisation. However, a complete tryptic digest of the same proteome will result in a large number of peptides including very hydrophilic ones. Fig. 2 shows the frequency distribution of hydrophilicities of tryptic peptides (derived

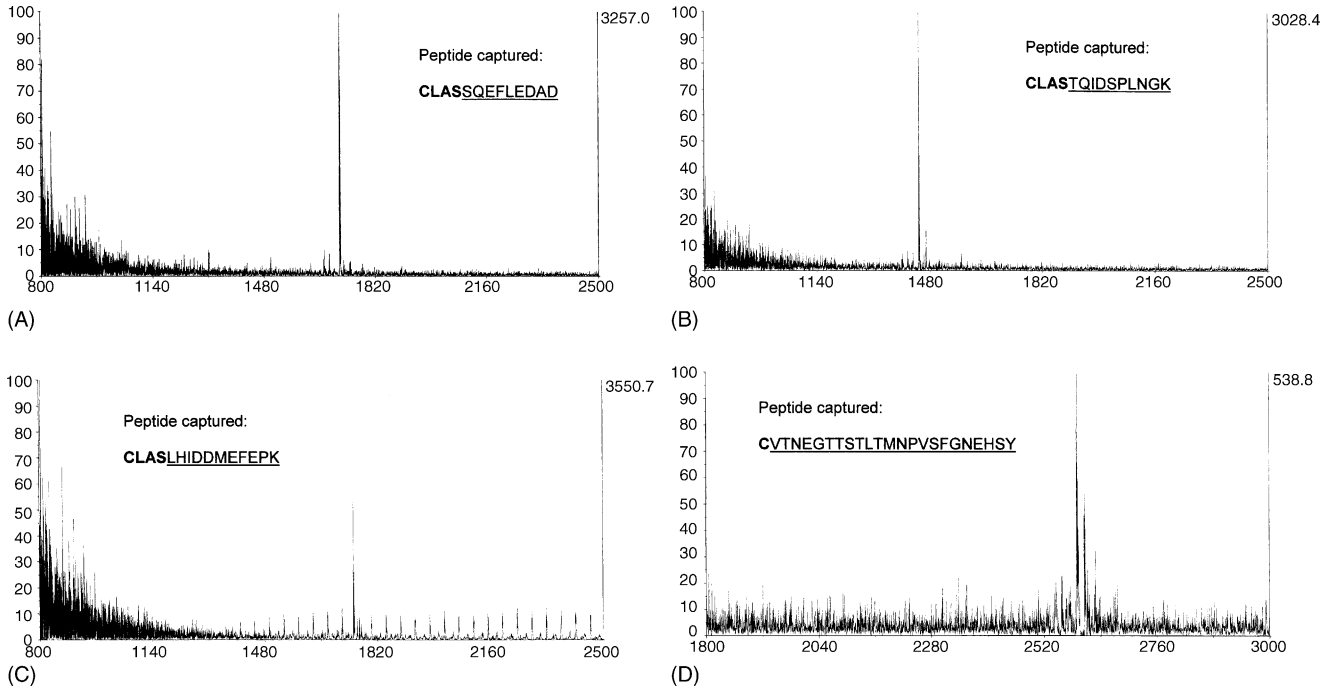


Fig. 1. MALDI-TOF MS traces are from Scrivener et al. [22] and were obtained from hydrogel chips following incubation of immobilised antibodies (on hydrogel pads) with a mixture of synthetic VCAM peptides (sequences shown). Tryptic digestion of the VCAM results in nearly 70 fragments, of which nearly 50 are hydrophilic. From these, five peptides were chosen for anti-peptide antibody development (single chain Fv(s), Table 1). Capture of the peptides indicated by MS traces for (A) CLASSQEFLEDADRK (the most hydrophilic of the VCAM peptides used, see Table 1); (B) CLASTQIDSPLNGK; (C) CLASLHIDDMEFEPK and (D) CVTNEGTTSTLTMNPVSFGNEHSY (lease hydrophilic of the VCAM peptides used). Underlined amino acids represent the tryptic fragments of the VCAM molecule used in scFv generation (few amino acids were added to each peptide to allow for effective conjugation). Peptide detection depends on the degree of their hydrophilicity. Images are reproduced with kind permission of the publisher (WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim).

Table 1
Vascular cell adhesion molecule (VCAM): predicted tryptic peptides*

Peptide sequence detection	Amino acid length	Hydrophilicity**	scFv(s) developed	Peptide
slenk	5	1.58		
SQEFLEDADR* **	**10**	**1.44**	**2 scFv tested**	**strong**** **
iettpesr	8	1.35		
sletk	5	1.02		
TQIDSPLNGK	**10**	**0.96**	**2 scFv tested**	**strong**
gdhlmk	6	0.88		
LHIDDMEFEPK	**11**	**0.88**	**2 scFv tested**	**weak**
dagvyecesk	10	0.84		
tqidsplsgk	10	0.69		
dpeihlsgpleagk	14	0.61		
getiledfledtdmk	17	0.52		
qstqtlyvnvapr	13	0.47		
gsyslveaqk	10	0.44		
vtnegttstltmnpvsfgnehsylctatcesr*** **	**32**	**0.41**		
VTNEGTTSTLTMNPVSFGNEHSY	**23**	**0.58**	**3 scFv tested**	**very weak**
medsgvylceginqagr	17	0.39		
segtnstltlspvsfenehsylctvtcghk	30	0.38		
aetgdtvlk	9	0.32		
qlpngelqplsenatltlistk	22	0.31		
pftveispgpr	11	0.29		
ntvisvnpstk	11	0.24		
elqvyispk	9	0.23		
lhidemdsvptvr	13	0.21		
vgsqlr	6	0.20		
sltldvqgr	9	0.18		
evelivqek	9	0.13		
vpsvypldr	9	0.12		
slemtfiptiedtgk	15	0.12		
lqeggsvtmtcsseglpapeifwsk	**25**	**0.10**		
SSEGLPAPEIFWSK	**14**	**0.35**	**3 scFv tested**	**not detected**
sidgaytir	9	0.04		
giqvelysfpr	11	0.03		

* Five amino acids or longer peptides shown only
** Kyte and Doolittle hydrophilicity predicted as in [14]
*** Underlined (dotted line) are the peptides used to develop anti-peptide affinity reagents (scFv single chains) using phage display approach (as in [16])
**** Peptide precipitaiton by scFv fragments as detected by MALDI-MS directly from silicon chip surface (as in [22]). See also Figure 1.
***** Only partial tryptic peptide sequence (underlined) was used for peptide synthesis and scFv generation. Syntetic peptide (sequence and predicted hydrophilicity) shown underneath

the same human proteome) and calculated for the single most hydrophilic peptides from each of the human proteins with the lower size limit of 6, 7 and 8 amino acids. The mode (i.e. the most frequently occurring hydrophilicity value in a distribution) for six amino acids or longer peptides is 2.25, for seven amino acids and longer peptides 1.8 and for eight amino acids or longer tryptic peptides 1.57. All these values are much higher than the 0.45 mode value obtained for full-length proteins or 0.58 the lowest peptide hydrophilicity yielding functional antibodies in the report by Scrivener et al. [22]. Overall, tryptic digestion results in that 98.7% of all human proteins will have their most hydrophilic peptides in the range 0.45–4.5, i.e. tryptic digestion will yield sufficiently hydrophilic peptides from nearly all (even very hydrophobic) human proteins. Therefore, the affinity peptidomics approach facilitates antibody generation through both better (more hydrophilic and immunogenic) peptide sequences and the availability of antigens (synthetic peptides). There is an added advantage to this approach in that the assay proteins now become small peptide fragments that are predictable in silico on the basis of sequence alone (protein or nucleotide). The reagents for affinity molecule selections are more easily generated via peptide synthesisers. The affinity peptidomics approach suits well both monoclonal (or recombinant) and polyclonal antibodies. Scrivener et al [22] has shown that polyclonal antibody can also be characterised for their antigen specificity (epitope mapping using peptidomics approach). Anti-protein (anti-human serum albumin HSA) polyclonal antibodies were immobilised on small hydrogel pads on a silicon chip, incubated with tryptically-digested human albumin and analysed by MALDI MS. The spectrum (Fig. 3) revealed three HSA peptides, representing the epitopes recognised by the polyclonal antibody preparation used. This approach can miss epitopes cut through by trypsin,

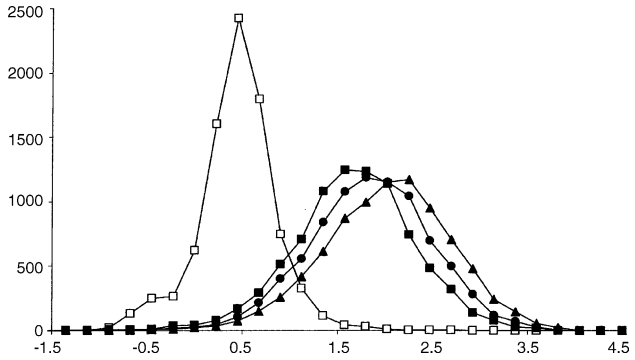


Fig. 2. Frequency distribution of hydrophilicities. Hydrophilicity values were calculated for human proteins (open boxes, the mode is 0.45, prediction based on Swiss Prot 40.27 and Kyte and Doolittle weights as in [14]) and for their tryptic peptides. Frequency of distribution of peptide hydrophilicites (for the most hydrophilic peptides—one for each human protein) are plotted: for peptides with the minimum length of six amino acids (triangles), minimum length of seven amino acids (diamonds) and eight amino acids (squares). The mode for six amino acids or longer peptides is 2.25, for seven amino acids and longer peptides −1.8 and for eight amino acids or longer tryptic peptides −1.57.

but these can be detected if another protease is used in a parallel experiment. An important feature of peptidomics is its compatibility with mass spectrometry (eliminating the need for protein/peptide labelling) and with protein microarrays, which can provide a fast route to proteome-scale analyses (Fig. 4A).

One of the major drawbacks of any affinity assay-based technique, including peptidomics, is the availability and the cost of traditional antibodies (capture reagents). Unlike nucleic acids, which are both information carriers and perfect affinity ligands, every protein or peptide requires the production of its own unique affinity reagent (e.g. an antibody) the development of which, unlike the synthesis of an oligonucleotide or purification of a PCR product, may require significant amounts of time and resources. Another specific difficulty is generally lower affinities of anti-peptide antibodies (or other affinity reagents). However, the much wider choice of peptides (i.e. epitopes) becoming available following a proteolytic digestion (as discussed above) almost certainly allows for a better choice (of an antigen) to be made for antibody generation.

5. Combinatorial peptidomics—peptide mixture refinement through selective depletion or enrichment

Combinatorial peptidomics allows the composition of a protein mixture to be determined by assaying peptides directly from crude proteolytic digests *without using antibodies or any other affinity selection* [23,24]. The successful implementation of a combinatorial peptidomics strategy requires methods for the selective removal and, if possible recovery, of subsets of peptides from the entire mixture resulting from proteolytic digestion. A general approach

Table 2
The amino acid side chain groups which can potentially be targeted as depletion or enrichment 'handles'

Amino acid	Side chain	pK_a	Potential reactivity
Arg	Guanidino	12.0	Dicarbonyl condensation Oxidation
Cys	Thiol	9.0–9.5	Disulphide formation Metal complexation
His	Imidazole	6.0–7.0; 14.4	
Lys[a]	Alkylamino	10.4–11.1	Acylation, alkylation Arylation, amidination
Met	Sulphide		oxidation
Trp	Indole		Electrophilic substn
Tyr	Phenol	10.0–10.3	Weak nucleophile Electrophilic substn
Phospho-Tyr	Aryl phosphate	∼1.6, ∼6.6	Esterification

[a] In a mixture of peptides generated by digestion with trypsin Lys cannot serve as a target for selective depletion but it is a candidate in other proteolytic mixtures.

to mixture refinement can in principle utilise the selective chemical reactivities of the side chains of individual amino acids, on which there is a considerable body of accumulated research [25–29]. Combinatorial peptidomics utilises the original peptidomics approach, in which protein samples are proteolytically digested using one or a combination of proteases [21,22]. However, in place of affinity purification, the peptide pool is depleted through chemical cross-linking of a subset of peptides to a solid support. Any of the six chemically reactive amino acid side chains (sulfhydryl groups of cysteines, thioether groups of methionines, imidazolyl groups of histidines, guanidinyl groups of arginines, phenolic groups of tyrosines and indolyl groups of tryptophans) can potentially be used to deplete (or enrich) a sample of the peptides, which contain them, in a specific and fully predictable *amino acid content-dependent* manner. These are listed in Table 2, together with some indication of reactivity.

Sequential chemical depletion (i.e. application of differently selective absorbents) will yield an increasingly depleted and therefore simplified mixture making it compatible with direct MS detection. The amino acid filtering (depletion) step may be repeated using combinations of up to six such filters (equivalent to a six-dimensional separation) or until the complexity of the peptide pool and the amino acid complexity of the remaining peptides is decreased to a desired level, suitable for mass spectrometric detection. As an estimate of whether full combinatorial depletion (or enrichment, or their combination) would be sufficient to deplete a "typical" human cell proteome (i.e. expressing tens of thousands of proteins and hence releasing millions of peptides following a tryptic digestion) one can assume that if every amino acid occurs once every 20-mer long amino acid fragment and each tryptic peptide will have on average 10 amino acids, then using only two different amino acid depletion steps, all such tryptic peptides can be precipitated (using for example four filters may precipitate most of 5-mer peptides and all six filters may precipitate most of the three amino acid long fragments). This is of course a very rough estimate, but it illustrates that the range of

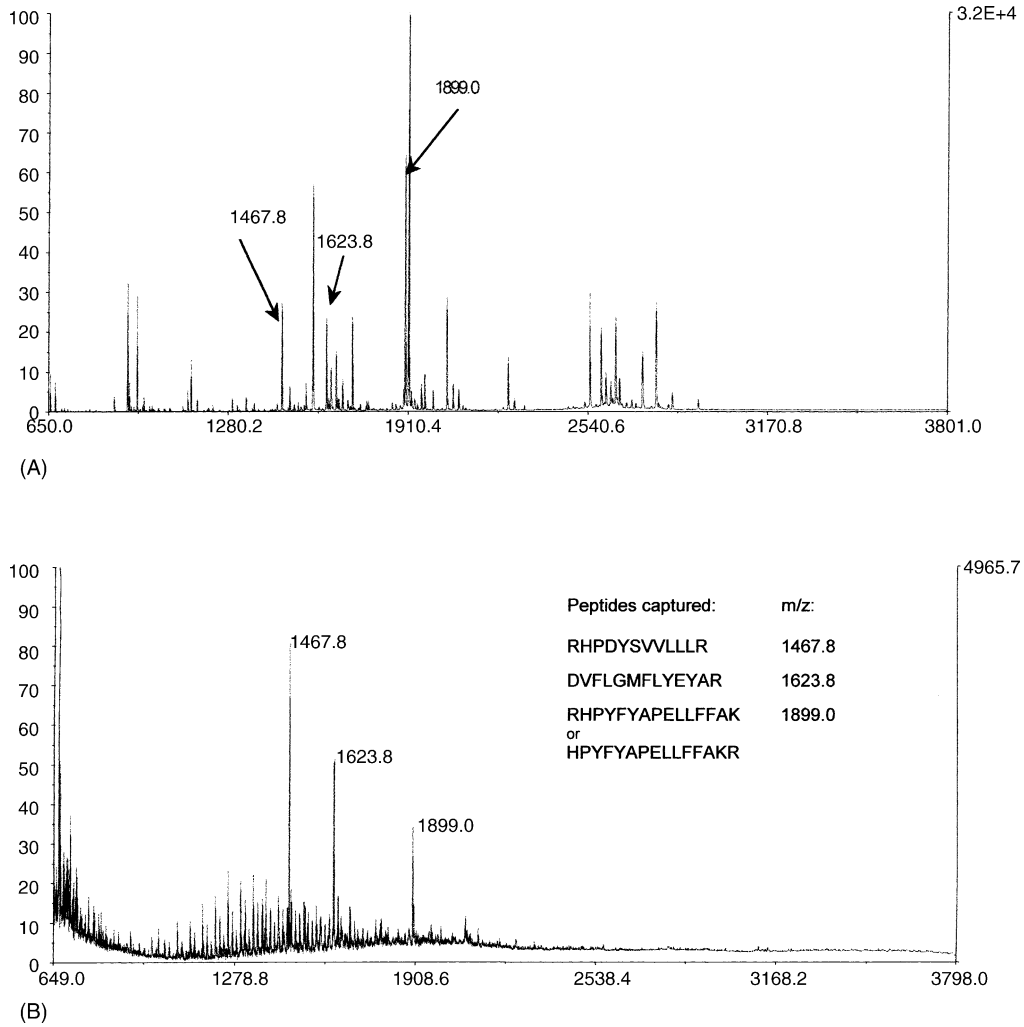


Fig. 3. MALDI-TOF MS traces of tryptic HSA peptides (from Scrivener et al. [22]). (A), prior to incubation with chips of immobilised antibodies. (B), A MALDI-TOF MS spectrum obtained from a gel pad following immobilisation with polyclonal antialbumin and incubation with a tryptic digest of HSA. The three major peaks observed are tryptic fragments of HSA as predicted by in silico digestion (seqeunces shown on panel (B)). The same peaks are highlighted in the spectrum of the original tryptic digest (panel (A)). MS traces are form Scrivener et al. [22] and reproduced with kind permission of the publisher (WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim).

peptide lengths covered and degree of depletion can be chosen to suit a typical mass spectrometric detection (100–1000 peptides per sample, with peptide lengths over 5–6 amino acids. It is important to note here, that because only those peptides that do not contain an amino acid recognised by the amino acid filter(s) remain in the mixture, the depleted peptide pools will contain peptides of reduced amino acid compositional complexity, which further facilitates the analysis of mass spectra and permits a greater number of peptide peaks to be identified from a single mass spectrum (Fig. 4B).

Many reagents have been described and studied but selectivity is not absolute and remains to be refined especially at the peptide, as distinct from protein, level. The practical implementation of the approach will require the formulation of this chemistry in a solid-phase format; in this way reactive peptides only will be retained on the solid-phase and the eluate will be a depleted mixture. If the retention chemistry can be reversed then clearly a mixture characterised by the presence of one particular amino acid can be obtained. A complete strategy (e.g. the optimum sequence of absorbents) will require a detailed appreciation of the reaction conditions appropriate to the individual amino acid side chain chemistries.

The chemical reactivities of the individual amino acid side chains and especially the identities and reaction conditions of selective reagents are presented in the following sections, which also describe the current situation regarding implementation in a solid phase procedure and reversibility of side chain reaction.

5.1. Arginine

Arginine (Fig. 5A) condensation reaction (Table 3) has not been reported in the solid phase mode. Otherwise, the reaction

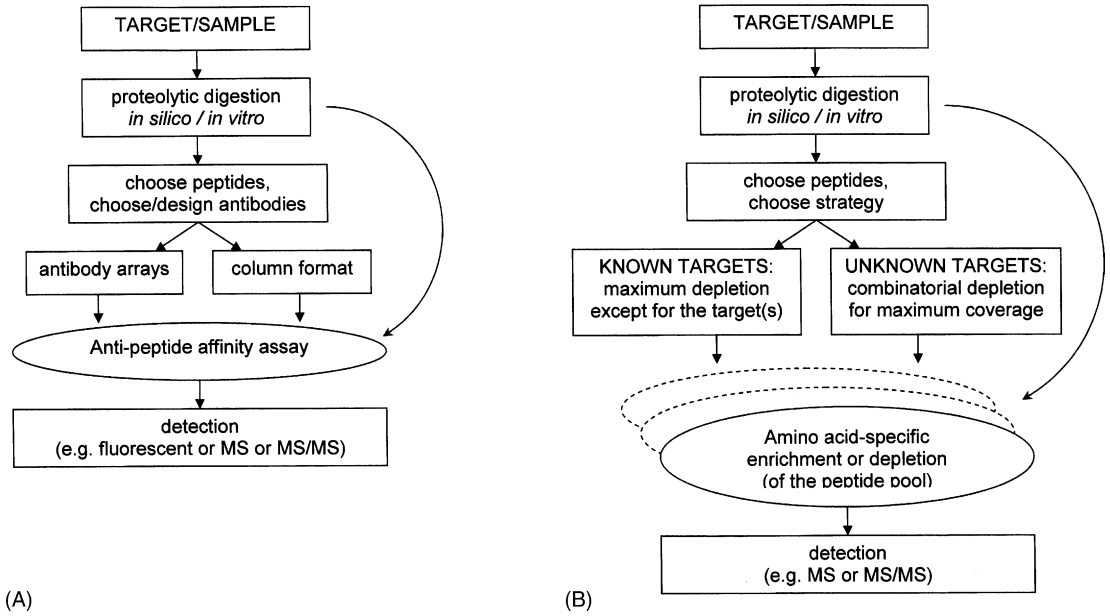


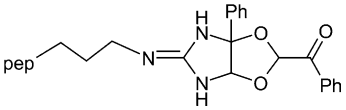
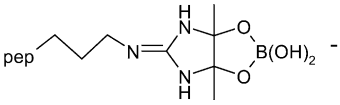
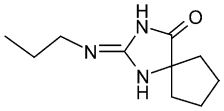
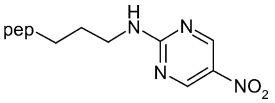
Fig. 4. A scheme illustrating affinity peptidomics (A) and combinatorial peptidomics (B). Affinity peptidomics relies on antibodies to capture proteolytic peptides. In combinatorial peptidomics, peptides are selected by depleting (or enriching) crude proteolytic digests by chemical cross-linking of the peptides to solid support (e.g. beads). Peptidomics can be applied to protein identification, quantitation, expression profiling, antibody characterisation and epitope mapping. It is multi-platform compatible and capable of simultaneously analysing proteins which differ strongly in their physico-chemical properties.

is carried out at room temperature in pH 7.5–8.0 buffers of which borate [30] and particularly bicarbonate [31] can have a specific accelerating effect. In studies on intact proteins, only partial modification of the total number of arginines has been observed, with those in a hydrophobic environment favoured. Arginine modification can be accompanied by side reactions at α- and Lys-ε-amino groups and, in the presence of oxygen under UV irradiation, at other residues, especially tryptophan and histidine [32]. Some classical reagents for selective modification of arginine are shown in Fig. 5B and Table 3.

5.2. Cysteine

The thiol group of cysteine (Fig. 5C) is the most reactive amino acid side chain and has been used as a handle for trapping and recovery of cysteine-containing peptides. Disulphide bonds between cysteinyl residues in native proteins will have been reduced to the thiol form prior to proteolytic digestion [33]. Cysteine-selective reagents, which could in principle be modified to solid-phase versions are listed in Table 4. Gaevert et al. [34] have reported a procedure for the isolation of cysteine-containing peptides after derivatisation at

Table 3
The classical reagents for selective modification of arginine

Reagents	Product	Comment	References
Phenylglyoxal		Selective for Arg in borate /bicarbonate buffers	[31,44]
Butane-2,3-dione in borate pH7.5		Reversed by hydroxylamine at pH 7	[30,45,46]
Cyclohexane-1,2-dione at pH 12			[50]
Nitro-malondialdehyde at pH 12–14		Irreversible	[51]

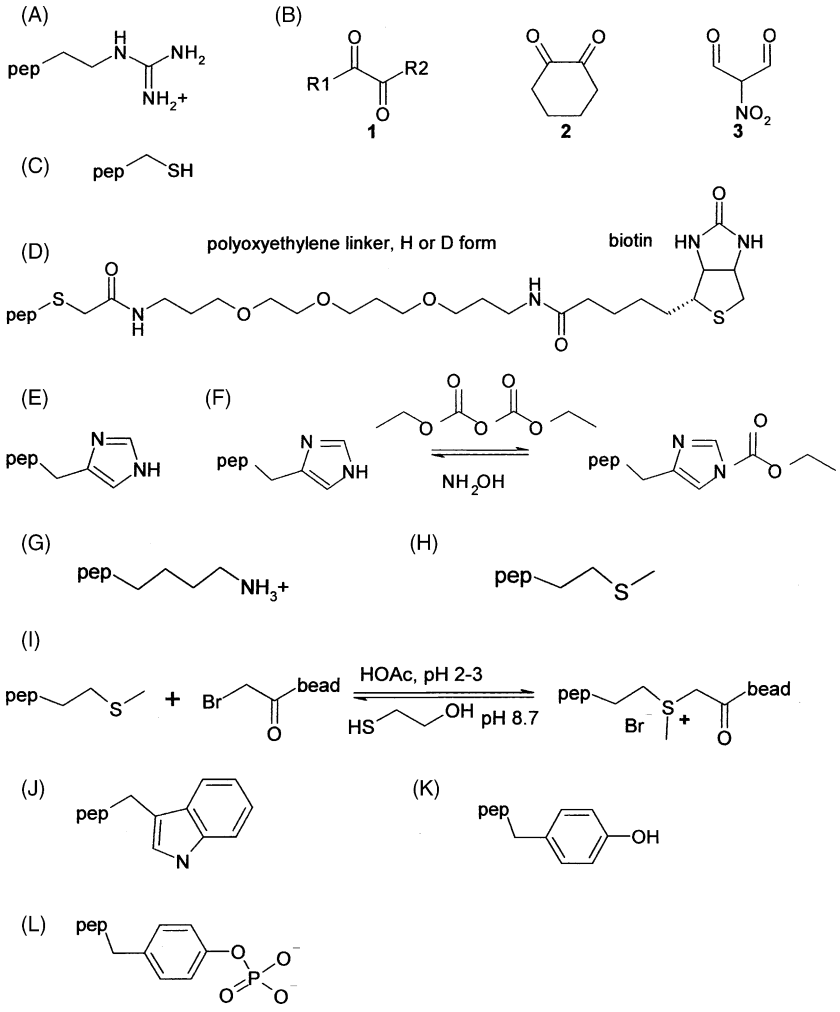


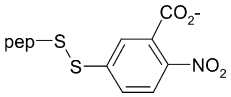
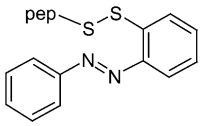
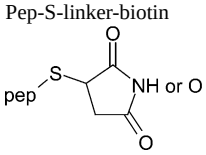
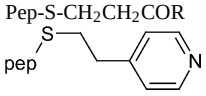
Fig. 5. Reactive amino acids and selective reagents for their modification. (A) Arginine. (B) The classical reagents for selective modification of arginine are dicarbonyl compounds of which phenylglyoxal (1, R1 = phenyl, R2 = H) [31,44], 2,3-butane-2,3-dione (1, R1 = R2 = CH_3) [30,45, 46], glyoxal (1, R1 = R2 = H) [47], 4-hydroxyphenylglyoxal (1, R1 = 4-OHPh) [48,49], 1,2-cyclohexanedione (2) [50] and nitro-malondialdehyde (3) [51] are the most studied. (C) Cysteine. (D) Commercially-available cysteine derivativatising reagent (4) based on iodoacetamide linked to a biotin terminus. (E) Histidine. (F) Diethyl pyrocarbonate is the classical selective reagent for modification of histidine. The reaction is reversed by hydroxylamine at neutral pH. (G) Lysine. (H) Methionine. (I) Methionine specific beads bearing bromoethanoyl groups, available commercially in a solid phase format [42]. (J) Tryptophan. (K) Tyrosine. (L) Phosphotyrosine.

the protein level with Ellman's reagent. After digestion with trypsin the hydrophobic peptides were isolated by reverse-phase high-performance liquid chromatography, regenerated by reduction and chromatographed again. The method was applied to proteomes of human platelets and enriched human plasma. A significant number of low abundance proteins were thereby identified in addition to extremely abundant ones.

Reversible covalent thiol-disulphide exchange chemistry has been exploited in a solid phase mode by Wang and Regnier [35] for the selection of cysteine-containing peptides from a tryptic digest of *Escherichia coli* cell lysate. Lysate was first reduced with dithiothreitol, trapped on a column of reversed phase silica and labelled at thiol groups in situ with 2,2′-dipyridyl disulfide at pH 5. Labelled proteins were then eluted from the column and digested with trypsin under denaturing conditions. After exchange into 0.1 M acetate buffer (pH 5.0) cysteine-containing peptides were captured by disulphide interchange on a column of reduced thiopropyl Sepharose gel. Binding could be monitored by the release of thiopyridone (λ_{max} = 343 nm) during this process. After washing with pH 5 and 7 buffers, cysteine-containing peptides were released from the gel with dithiothreitol containing EDTA. The composition of the peptide digest was thereby considerably simplified prior to analysis by chromatography and MALDI MS. Reversible disulphide exchange chemistry using 2-mercaptopyridine immobilised on a column has also been described [36].

Isotope-coded affinity tags (ICAT) reagents [37] are commercially available derivativatising agents (Fig. 5D) based on iodoacetamide linked to a biotin terminus. After the reaction of denatured protein in solution and tryptic digestion the biotin moiety is used as a specific affinity ligand

Table 4
The classical reagents for selective modification of cysteine

Reagents	Product	Comment	References
Aromatic disulphide e.g. Ellman's reagent 2-mercaptopyridine		Selective and reversible	[52] [36]
Azobenzene-2-sulphenyl bromide		Reversible	[53]
Iodoacetate, iodoacetamide	Pep-S-$CH_2CO_2^-$	Histidine may react	[54,55]
ICAT reagents	Pep-S-CH_2CONH_2		
	Pep-S-linker-biotin	Biotin handle	[37]
N-ethylmaleimide, maleic anhydride		Amino groups may react	[56,57]
Ethyleneimine	Pep-S-$CH_2CH_2NH_3$+		[58]
Acryloyl compounds	Pep-S-CH_2CH_2COR	SH more rapidly than NH_2 at pH7	[59,60]
4-Vinylpyridine			[61]
(2-Bromoethyl) trimethylammonium bromide	Pep-S-$CH_2CH_2N^+(Me)_3$	α-Amino, Met may react	[62]
Methyl-4-nitrobenzene sulphonate	Pep-S-CH_3	Highly selective for Cys	[63]
4-Mercuribenzoic acid	Pep-S-Hg-$PhCO_2^-$		[64,65]

for selectively collecting derivatised peptides on an avidin column. Cysteine-containing peptides can be subsequently desorbed with formic acid and analysed by LC–MS/MS. A cleavable version of the ICAT reagent, which simplifies the mass spectra of the isolated peptides, is now available.

5.3. Histidine

The NH group of histidine (Fig. 5E) can be readily derivatised by acylating agents but thiol and other amino groups are generally more reactive in this respect and acyl histidines are labile to hydrolysis. The classical selective reagent is diethyl pyrocarbonate (Fig. 5F), which reacts selectively at acidic pH [38]. The reaction is rapidly reversed by hydroxylamine at neutral pH. No solid phase version of this chemistry has been reported.

Regnier and co-workers [39] have developed a scheme for the selection of histidine-containing peptides by exploiting the metal-complexation property of the imidazole side chain. Peptides were captured on a TSK gel Chelate-5PW column in the Cu^{2+} form at pH 7.5 and released by washing with buffer at pH3.9. In order to achieve selectivity towards histidine, peptides were first acylated at primary amino groups with succinic anhydride. By the sequential application of cysteine- and histidine-selective absorbent columns Regnier and co-workers isolated two peptides from an *E. coli* lysate digest, one of which was identified unequivocally from 40 possible peptides of the same mass and the other narrowed down to four of a possible 45 in the digest. The efficiencies of different histidine-selective resins have recently been compared [40].

5.4. Lysine

The nucleophilic amino group of lysine (Fig. 5G) readily reacts with a variety of reagents shown in Table 5. For example, a 1,3-dioxobutanyl derivative of polystyrene, suitable for use with combinatorial peptidomics experiment, is currently available from Sigma-Aldrich (Cat. no. 55,147-3).

5.5. Methionine

The thioether group of methionine (Fig. 5H) reacts by oxidation or alkylation as shown in Table 6. At acidic pH alkylation proceeds more rapidly than with cysteine and amino groups (which are fully protonated). This chemistry has been realised in a solid phase format [24,41] with commercially available beads [42] bearing bromoethanoyl groups (Fig. 5I). Soloviev et al. [24] reacted a mixture of ten synthetic peptides, five of which contained methionine in different positions along their sequences, at 22 °C for 18 hours. Subsequent analysis of the supernatant showed no Met-containing peptides. In another application of this absorbent, McEldoon et al. [41] recovered five peptides from an apomyoglobin/apoferritin digest of which three had masses (MALDI) which matched expected Met-containing sequences. However, the peptides

Table 5
The classical reagents for selective modification of lysine

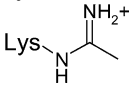
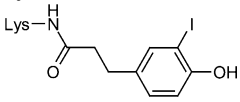
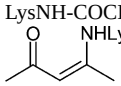
Reagents	Product	Comment	References
Trinitrobenzene sulphonates	LysNH-Ph(NO_2)$_3$	α-Amino and SH groups compete	[66]
Acid anhydrides e.g., ethanoic, succinic maleic, citraconic	LysNH-COR	Other side chains react, esp α-amino; reversible in acid	[67,68]
Ethyl acetimidate			[69]
Aldehydes RCHO e.g. pyridoxal phosphate	LysN=CHR	Reversible by hydrolysis	[70]
N-(OH)-succinimido esters e.g. (Bolton-Hunter reagent)		Selective for Lys at pH 8.5, α-amino at pH 6.5	[71]
Diketene	LysNH-$COCH_2OCOCH_3$	Tyr and Ser derivs reversible	[72]
2,4-Pentanedione		Arg reacts but more slowly at pH 7	[73]

Table 6
The classical reagents for selective modification of methionine

Reagents	Product	Comment	References
Halo-acids and amides	pep$CH_2S^+(CH_3)COR$	Selective in acid	[74]
β-Propiolactone	pep$CH_2S^+(CH_3)CH_2CH_2CO_2H$		[75]
Alkyl halides	pepCH_2 $S^+(CH_3)R$		[76]

appeared in both native and partially oxidised forms. In the same report, 34 peptides and proteins were identified from a tryptic digest of an *E. coli* lysate.

5.6. Tryptophan

The most selective compounds for modification of tryptophan residues (Fig. 5J) are the electrophilic reagents based on 2-hydroxy-5-nitrobenzyl bromide (Koshland's reagent) and 2-nitrobenzenesulfenyl chloride (Table 7). Peptide capture by reversible disulphide derivatisation of tryptophan side chains in the solid phase mode has been reported by the Biomolecular Technologies group [41]. Peptides with the expected masses (including an additional 32 units) were recovered with high selectivity from tryptic digests of lysozyme and an apomyoglobin/apoferritin mixture.

5.7. Tyrosine

The common modifications of tyrosine (Fig. 5K), such as nitration and iodination have little potential for solid phase trapping of peptides. Reactions with *N*-acetylimidazole and diazonium compounds have been widely studied but are not completely specific. A diazonium functionalised resin based on polystyrene is available and can be applied for combinatorial peptidomic applications (Table 8).

Table 7
The classical reagents for selective modification of tryptophan

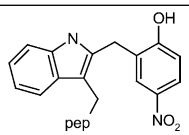
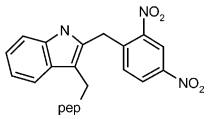
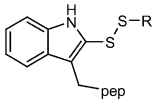
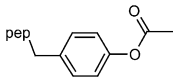
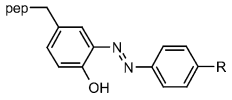
Reagents	Product	Comment	References
2-Hydroxy-5-nitrobenzyl bromide		Gives mixtures of products	[77]
2,4-Dinitrobenzene sulfenyl chloride		Cysteine reacts	[78]
Chlorodisulphanyl derivatives		Mercaptoethanol releases the 2-thiol derivative	[41]

Table 8
The classical reagents for selective modification of tyrosine

Reagents	Product	Comment	References
N-acetylimidazole at pH 7.0–7.5		Reversible with hydroxylamine at pH 7.5	[79]
Diazonium compounds			[80]

5.8. Phosphotyrosine

The phosphate group of phosphotyrosine (Fig. 5L) can be reversibly condensed with amines (making it suitable for both depletion and enrichment combinatorial peptidomics modes) but carboxyl groups must first be protected. The Biomolecular Technologies group has reported a procedure for solid phase capture and release of phosphotyrosine-containing peptides (after methyl esterification) on an imidazole-functionalised column [41]. Phospho-serine and threonine-containing peptides were not released under the conditions used.

Ficcaro et al. [43] have described a procedure for the selective isolation of tryptic phosphopeptides by affinity chromatography. Peptides were converted to their methyl esters, fractionated by immobilized metal-affinity chromatography and analyzed by nanoflow HPLC/electrospray ionization mass spectrometry. More than 1000 phosphopeptides derived from 216 sequences defining 383 sites of phosphorylation were determined in the digest of a whole-cell lysate from saccharomyces cerevisiae. The approach by Ficcaro et al. is similar to the combinatorial peptidomics approach in targeting individual amino acids (which are phosphorylated), but it does not rely on chemical crosslinking of phosphorylated peptides and thus provides for less quantitative analysis.

6. Conclusion

The original combinatorial peptidomics approach relies on the chemical reactivities of amino acids and therefore the amino acid content of the peptides (i.e. their information content), rather than their physical properties. This makes combinatorial peptidomics very similar in principle to nucleic acid assays (Southern blots, Northern blots, microarrays) where analysis is based on the nucleic acid sequence, i.e. the *information content*, and offers a further advantage over existing proteomics techniques including affinity selection-based approaches such as affinity peptidomics (limited by the availability of anti-peptide antibodies), multidimensional chromatography or immobilised metal-affinity chromatography. With more suitable resins becoming available commercially, the emphasis now is on the development of more selective sorbents and on optimising and streamlining the depletion or enrichment procedures to allow for a faster analysis. Both peptidomics approaches described allow for a relative quantitation to be performed easily. However, the problem of absolute quantitation of individual proteins/peptides in a sample may require the use of isotopically labelled references.

Proteomics is of very high value in functional genomics analysis, but it currently represents a significant bottleneck to progress in this area. We do not believe that further refinements of 60-year old 2D gel system (which has since proven to have serious technical and economic limitations) or stacking yet another chromatography separation step on top of existing 2D or 3D separations, can resolve the challenges facing modern proteomics.

Peptidomics and combinatorial peptidomics were conceived at the turn of the 21st century by a group of dedicated scientists working for then the world leading industrial biotechnology company, Oxford Glycosciences, one of the first spin-off companies from the University of Oxford. These new technologies are the culmination of 14 years of protein research and technology development mastered by Oxford Glycosicences.

Peptidomics and combinatorial peptidomics are designed to supersede the 2D and LC–MS approaches and define the enabling new technology, capable of quick identification *and* quantitation of many hundreds of proteins in a straightforward assay. Peptidomics compares to traditional proteomics techniques just as digital signal processing compares to analogue systems—*it forms the basis of the high-throughput proteomics technologies of the future.*

References

[1] J.D. Watson, F.H. Crick, Nature 171 (1953) 737.
[2] F. Sanger, S. Nicklen, A.R. Coulson, Proc. Natl. Acad. Sci. U.S.A. 74 (1977) 5463.
[3] E.S. Lander, L.M. Linton, B. Birren, C. Nusbaum, et al., Nature 409 (2001) 860.
[4] J.D. McPherson, M. Marra, L. Hillier, R.H. Waterston, et al., Nature 409 (2001) 934.
[5] B.B. Haab, M.J. Dunham, P.O. Brown, Genome Biol. 2 (2001), research 0004.
[6] R. Barry, T. Diggle, J. Terrett, M. Soloviev, J. Biomol. Screening 8 (2003) 257.
[7] R. Barry, M. Soloviev, Proteomics, in press.
[8] S.D. Patterson, R. Aebersold, Electrophoresis 16 (1995) 1791.
[9] K. Gevaert, M. Goethals, L. Martens, J. Van Damme, A. Staes, G.R. Thomas, J. Vandekerckhove, Nature Biotechnol. 21 (2003) 566.

[10] M.A. Cooper, Anal. Bioanal. Chem. 377 (2003) 834.
[11] D.G. Myszka, Anal. Biochem. 15 (2004) 316.
[12] J.V. Olsen, S.E. Ong, M. Mann, Mol. Cell. Proteomics 3 (2004) 608.
[13] L.C. Mahan, L.D. McVittie, E.M. Smyk-Randall, H. Nakata, F.J. Monsma Jr., C.R. Gerfen, D.R. Sibley, Mol. Pharmacol. 40 (1991) 1.
[14] J. Kyte, R.F. Doolittle, J. Mol. Biol. 157 (1982) 105.
[15] T.P. Hopp, Pept. Res. 6 (1993) 183.
[16] V.F. de la Cruz, A.A. Lal, T.F. McCutchan, J. Biol. Chem. 263 (1988) 4318.
[17] J. Hanes, A. Pluckthun, Proc. Natl. Acad. Sci. U.S.A. 94 (1997) 4937.
[18] D.S. Wilson, A.D. Keefe, J.W. Szostak, Proc. Natl. Acad. Sci. U.S.A. 98 (2001) 3750.
[19] G. Fletcher, S. Mason, J. Terrett, M. Soloviev, J. Nanobiotechnol. 1 (2003) 1.
[20] M. Soloviev, Curr. Drug Discovery (2004) 25.
[21] R. Barry, A.E. Platt, E. Scrivener, M. Soloviev, J.Terrett, WO 02/25287.
[22] E. Scrivener, R. Barry, A. Platt, R. Calvert, G. Masih, P. Hextall, M. Soloviev, J. Terrett, Proteomics 3 (2003) 122.
[23] M.Soloviev, J.Terrett, WO 03/027681.
[24] M. Soloviev, R. Barry, E. Scrivener, J. Terrett. 1 (2003) 4.
[25] E. Gary, Chemical Modification of Proteins, Means and Robert E. Feeney, San Francisco, Holden-Day, 1971.
[26] T.S. Work, E. Work (Eds.), Laboratory Techniques in Biochemistry and Molecular Biology, North Holland Pub. Co., Amsterdam, 1975.
[27] H. Neurath, R.L. Hill, C.-L. Boeder (Eds.), The Proteins, 3rd ed., Vol II, Academic, NY, 1976 pp. 1-103.
[28] R.L. Lundblad, C.M. Noyes, Chemical Reagents for Protein Modification, vols. 1 and 2, CRC Press, Boca Raton, FL, 1984.
[29] N.C. Price (Ed.), Proteins Labfax, BIOS, Oxford, 1996, p. 287.
[30] J.F. Riordan, Biochemistry 12 (1973) 3915.
[31] S.T. Cheung, M.L. Fonda, Biochem. Biophys. Res. Commun. 90 (1979) 940.
[32] H. Fliss, T. Viswanatha, Can. J. Biochem. 57 (1979) 1267.
[33] W.W. Cleland, Biochemistry 3 (1964) 480.
[34] K. Gaevert, B. Ghesquiere, A. Staes, L. Martens, J. VanDamme, G.R. Thomas, J. Vanderckerkove, Proteomics 4 (2004) 897.
[35] S. Wang, F.E. Regnier, J. Chromatogr. A 924 (2001) 345.
[36] N.C. Price (Ed.), Proteins Labfax, BIOS, Oxford, 1996, p. 65.
[37] S.P. Gygi, B. Rist, S.A. Gerber, F. Turecek, M.H. Gelb, R. Aebersold, Nature Biotechnol. 17 (1999) 994.
[38] E.W. Miles, Methods Enzymol. 47 (1977) 431.
[39] S. Wang, X. Zhang, F.E. Regnier, J. Chromatogr. A 949 (2002) 153.
[40] D. Ren, N.A. Penner, B.E. Slentz, H.D. Inerowicz, M. Rybalko, F.E. Regnier, J. Chromatogr. A 1031 (2004) 87.
[41] http://www.bmtusa.com/pdf/asms02r1.pdf.
[42] http://www.bmtusa.com/methionine.html.
[43] S.B. Ficarro, M.L. McCleland, P.T. Stukenberg, D.J. Burke, M.M. Ross, J. Shabanowitz, D.F. Hunt, F.M. White, Nature Biotechnol. 20 (2002) 301.
[44] K. Takahashi, J. Biol. Chem. 243 (1968) 6171.
[45] J.A. Yankelov Jr., Biochemistry 9 (1970) 2433.
[46] J.A. Yankelov Jr., D. Acree, Biochem. Biophys. Res. Commun. 42 (1971) 886.
[47] J.D. Glass, M. Pelzig, Biochem. Biophys. Research Commun. 81 (1978) 527.
[48] C.L. Borders Jr., L.J. Pearson, A.E. Mclaughlin, M.E. Gustavson, J. Vasiloff, F.Y. An, D.J. Morgan, Biochim. Biophys. Acta 568 (1979) 491.
[49] R.B. Yamasaki, A. Vega, R.E. Feeney, Anal. Biochem. 109 (1980) 32.
[50] L. Patthy, E.L. Smith, J. Biol. Chem. 250 (1975) 557.
[51] A. Signor, G.M. Bonora, L. Biodi, D. Nisato, A. Marzotto, E. Scoffone, Biochemistry 10 (1971) 2748.
[52] G.L. Ellman, Arch. Biochem. Biophys. 82 (1959) 70.
[53] A. Fontana, F.M. Veronese, E. Scoffone, Biochemistry 7 (1968) 3901.
[54] A.M. Crestfield, S. Moore, W.H. Stein, J. Biol. Chem. 238 (1963) 622.
[55] J.D. Hempel, R. Pietruszko, J. Biol. Chem. 256 (1981) 10889.
[56] J.D. Gregory, J. Am. Chem. Soc. 77 (1955) 3922.
[57] D.G. Smyth, O.O. Blumenfeld, W. Konigsberg, Biochem. J. 91 (1964) 589.
[58] M.A. Raftery, R.D. Cole, J. Biol. Chem. 241 (1966) 3457.
[59] T.S. Seibles, L. Weil, Methods Enzymol. 11 (1967) 204.
[60] T.H. Plummer Jr., C.H.W. Hirs, J. Biol. Chem. 239 (1964) 2530.
[61] J.F. Cavins, M. Friedman, Anal. Biochem. 35 (1970) 489.
[62] H.A. Itano, E.A. Robinson, J. Biol. Chem. 247 (1972) 4819.
[63] R.L. Heinrickson, J. Biol. Chem. 246 (1971) 4090.
[64] P.D. Boyer, J. Am. Chem. Soc. 76 (1954) 4331.
[65] R. Benesch, R.E. Benesch, Methods Biochem. Anal. 10 (1962) 43.
[66] T. Okuyama, K. Satake, J. Biochem. (Tokyo) 47 (1960) 454.
[67] I.M. Klotz, Methods Enzymol. 11 (1967) 576.
[68] L. King, R.N. Perham, Biochemistry 10 (1971) 981.
[69] T. Sekiguchi, S. Oshiro, E.M. Goingo, Y. Nosoh, J. Biochem. (Tokyo) 85 (1979) 75.
[70] M. Rippa, C. Picco, S. Pontremoli, J. Biol. Chem. 245 (1970) 4977.
[71] G. Gaudriault, J.P. Vincent, Peptides 13 (1992) 1187.
[72] R.W. Mahley, K.H. Weisgraber, T.L. Innerarity, H.G. Windmueller, Proc. Natl. Acad. Sci. U.S.A. 76 (1979) 1746.
[73] H.B. Otwell, K.L. Cipollo, R.B. Dunlap, Biochim. Biophys. Acta 568 (1979) 297.
[74] H.G. Gundlach, S. Moore, W.H. Stein, J. Biol. Chem. 234 (1959) 1761.
[75] G.R. Stark, Adv. Protein Chem. 24 (1970) 261.
[76] T.P. Link, G.R. Stark, J. Biol. Chem. 243 (1968) 1082.
[77] H.R. Horton, D.E. Koshland Jr., Methods Enzymol. 25B (1972) 468.
[78] A. Fontana, E. Scoffone, Methods Enzymol. 25B (1972) 482.
[79] J.F. Riordan, B.L. Vallee, Methods Enzymol. 25 (1972) 500.
[80] J.F. Riordan, B.L. Vallee, Methods Enzymol. 25 (1972) 521.