JOURNAL OF MASS SPECTROMETRY
J. Mass Spectrom. 2001; **36**: 1249–1259

JMS

SPECIAL FEATURE: TUTORIAL

Mass spectrometry and combinatorial chemistry: a short outline

Antonio Triolo,[1]* Maria Altamura,[1] Franco Cardinali,[2] Alessandro Sisto[2] and Carlo Alberto Maggi[1]

[1]Menarini Ricerche SpA, via Sette Santi 3, 50131 Florence, Italy
[2]Menarini Ricerche SpA, via Tito Speri 10, 00040 Pomezia, Italy

Received 24 July 2001; Revised 31 August 2001; Accepted 19 September 2001; Published online 12 November 2001

The rapid evolution of combinatorial chemistry in recent years has led to a dramatic improvement in synthetic capabilities. The goal is to accelerate the discovery of molecules showing affinity against a target, such as an enzyme or a receptor, through the simultaneous synthesis of a great number of structurally diverse compounds. This is done by generating combinatorial libraries containing as many as hundreds or thousands of compounds. The need to test all these compounds led to the development of high-throughput screening (HTS) techniques, and also high-throughput analytical techniques capable of assessing the occurrence, structure and purity of the products. In order to be applied effectively to the characterization of combinatorial libraries, an analytical technique must be adequately sensitive (to analyse samples which are typically produced in nanomole amounts or less), fast, affordable and easy to automate (to minimize analysis time and operator intervention). Although no method alone can meet all the analytical challenges underlying this task, the recent progress in mass spectrometric (MS) instrumentation renders this technique an essential tool for scientists working in this area.

We describe here relevant aspects of the use of MS in combinatorial technologies, such as current methods of characterization, purification and screening of libraries. Some examples from our laboratory deal with the analysis of pooled oligomeric libraries containing $n \times 324$ (n = 1, 2) compounds, using both on-line high-performance liquid chromatography/MS with an ion trap mass spectrometer, and direct infusion into a triple quadrupole instrument. In the first approach, MS and product ion MS/MS with automatic selection of the precursor were performed in one run, allowing library confirmation and structural elucidation of unexpected by-products. The second approach used MS scans to characterize the entire library and also precursor ion and neutral loss scans to detect selectively components with given structural characteristics. **Copyright © 2001 John Wiley & Sons, Ltd.**

KEYWORDS: combinatorial chemistry; combinatorial libraries; electrospray ionization mass spectrometry; collision-induced dissociation; liquid chromatography

INTRODUCTION

The interest of pharmaceutical companies in employing combinatorial technologies in the process of drug discovery is driven by both economic and scientific reasons. From an economic point of view, the long time (typically, more than 10 years) and the increasing costs of introducing a new drug to the market led to the need to find lead compounds in a more cost- and time-effective manner than by the traditional approaches. The latter include the serial synthesis of new chemical entities or the random screening of collections of natural compounds from fermentation broths. In addition, present-day achievements in genetics and molecular biology are rapidly changing the scientific landscape: milestone events, such as the sequencing of the human genome, are generating a massive numbers of new targets involved in therapeutically relevant processes. All this has the consequence of dramatically increasing the need to be able to produce new chemical entities to test against those targets, and combinatorial chemistry is currently accepted as having the potential to satisfy these needs.

The large number of reviews that have been published since the early 1990s witness the general interest in this

*Correspondence to: A. Triolo, Menarini Ricerche SpA, via Sette Santi 3, 50131 Florence, Italy. E-mail: atriolo@menarini-ricerche.it

DOI: 10.1002/jms.238
Copyright © 2001 John Wiley & Sons, Ltd.

topic. Some recent reviews are listed in Ref. 1–3. During the process of lead-finding and optimization, as many as thousands of compounds are simultaneously generated using combinatorial techniques. Depending on numerous factors, such as the compound structures, the target type and the assay chosen to evaluate biological activity, the libraries can be prepared as mixtures (sometimes defined as 'pooled libraries') or as collections of single substances in separate vessels (sometimes defined as 'discrete libraries' or 'arrays'). The synthesized substances may exist free in solution or be supported on a solid phase. The most typical methodology for the preparation of pooled libraries is based on solid-phase synthesis, according to the split-pool technique pioneered by Furka *et al.*[4] for peptide libraries. Briefly, before the first reaction step, resin beads containing an initial substrate are split into multiple equal portions, which are separately made to react with each building block (e.g., each desired N-protected C-terminal amino acid in the case of a classical solid-phase peptide synthesis). After reaction, the resin beads are mixed and again split into the desired number of portions, which are reacted with the building blocks used for the second step, yielding all the possible combinations of the two series of building blocks. The split-pool process is repeated for the desired number of steps, producing a total number of compounds given by $N = a^x$, where a is the number of building blocks (considered constant for each step in this particular case) and x is the number of reaction steps in which a building block is introduced. Since each reaction is performed with a single building block on a separate portion of solid support, at the end of the synthesis a given bead contains only one compound. Depending on the resin loading capacity, each bead carries typically 100 pmol–5 nmol of compound. Collections of discrete compounds are prepared by parallel synthesis, either in the solid phase or in solution, often using highly automated instrumentation. Each compound is obtained in amounts from nanomoles up to millimoles.

To allow time- and cost-effective biological testing of the compounds prepared by combinatorial synthetic methods, high-throughput screening methodologies have been developed for the rapid evaluation of biological activity and are now well established, so that they are capable of screening up to 50 000 compounds per day, an order of magnitude more than the synthetic capabilities of current automated parallel synthesis instrumentation.[2] Besides high-throughput synthesis and high-throughput activity screening, a third issue of essential importance consists in the high-throughput analytical characterization of the combinatorial chemistry products (and eventually structural identification of the active components). This is currently considered the bottleneck of the entire process, and considerable effort is being devoted to the development of reliable high-throughput analytical methods in support of the synthetic and screening activities. General analytical problems, strategies and techniques have been discussed.[5–7] The tasks of analytical chemistry when faced with the analysis of libraries are not qualitatively different from those required by traditional serial synthetic methods, that is, basically, confirmation of the desired structure (e.g. by NMR, IR, MS, etc.) and purity evaluation (essentially, by chromatographic methods, elemental analysis, etc.) is desired. However, time constraints, the small amounts of each compound and the huge number of compounds synthesized in arrays or in pooled forms render complete characterization of each compound occurring in a library impractical or even impossible. Thus, analytical strategies have been developed to characterize with reasonable confidence combinatorially generated compounds. In this approach, emphasis is placed on the importance of analytical work in the development phase of the library: good characterization of standards, products and by-products and valid assistance in the set-up of reactions conditions during the synthesis of model libraries greatly enhance the confidence in the successful synthesis of the final library, even when the latter cannot be fully characterized.[5] In this preliminary phase, the methods and techniques do not differ greatly from those used during the synthesis of single compounds. For the analytical characterization of libraries, the methods used may vary depending on the type of library synthesized, but they should fulfil the following basic requirements:

- high degree of automation during the whole process of sampling, data acquisition, data reduction and reporting;
- capability of providing information about the occurrence, structure and purity of the sample in the shortest time as possible;
- sensitivity, in order to minimize sample consumption; and
- affordability and robustness, in order to analyse a large number of samples without loss in performance.

At present, there is no perfect technique which fully satisfies all these requirements, although considerable improvements have been reported in all the currently used analytical techniques, such as gel-phase,[8] magic angle spinning[9,10] and high-performance liquid chromatography (HPLC)/NMR.[11] There has been similar progress in new sampling methods including Fourier transform (FT) IR and FT-Raman spectroscopy.[12] However, the intrinsic sensitivity and specificity of mass spectrometry (MS), its inherent capability of analysing complicated mixtures (especially using high-resolution MS and tandem MS) and the recent progress in instrumentation render this technique an essential and versatile tool not only for the structural and analytical characterization of combinatorial libraries, but also for their high-throughput purification, biological screening and kinetic-metabolic profiling. The role of MS in combinatorial technologies has been described in several reviews.[7,13–16] It is to be understood, anyway, that the solution of very complex analytical problems arises from the combined, complementary use of the information coming from different techniques. This Tutorial article gives an outline of the MS methods currently used in the field of combinatorial technologies.

The successful use of MS in combinatiorial chemistry relies mainly on its sensitivity, speed and specificity. As stated above, low sample consumption is mandatory for the analytical characterization of compounds sometimes produced in amounts of a few hundred picomoles. Current LC/MS instrumentation can routinely perform this task with <1 pmol, analysing automatically hundreds of samples per hour. However, perhaps the most important advantage is

Copyright © 2001 John Wiley & Sons, Ltd.

that, in most cases, all this can be done without compromising significantly the wealth of information obtained. Exact mass measurements, for example, can currently be obtained with great sensitivity and speed,[17] and detailed structural information is obtained in real time by automatic precursor ion selection for 'on the fly' product ion spectra in commercial tandem mass spectrometers. Another advantage of MS is that it is a relatively universal technique: analytes do not need to bear a UV-absorbing chromophore or be fluorescent in order to be detected, and the availability of different ionization techniques is such that most organic molecules are amenable to MS analysis. This feature is also advantageous for combinatorial libraries, which are nowadays composed of very heterogeneous classes of compounds. In fact, although the first libraries were composed mostly of peptides or oligonucleotides, prepared by well-developed solid-phase methods, the inherent poor stability and bioavailability of these biopolymers led to the need for the synthesis of different classes of molecules, more metabolically stable and amenable to oral administration. It is to be expected that the tendency to discover and develop orally active compounds will continue in the future, giving rise to an ongoing interest in the synthesis of small organic molecules, mostly having molecular masses below 700–800 Da and medium-to-high polarities. Such compounds can be successfully analysed, at high throughput, using atmospheric pressure ionization (API) techniques such as electrospray ionization (ESI)[18] and atmospheric pressure chemical ionization (APCI).[19] In normal operation, the analytes usually appear as protonated (or deprotonated in the negative ion mode) molecular ions, yielding simply interpretable information on the molecular mass. Structure-indicating fragments can be obtained easily by increasing the energy of the collisions between ions and neutrals in the interface (in-source collision-induced dissociation (CID)), and also by tandem MS (MS/MS) when a suitable analyser (e.g. a triple quadrupole or an ion trap) is used. Another ionization technique used in the MS analysis of combinatorial libraries is matrix-assisted laser desorption/ionization (MALDI),[20] which is commonly combined with time-of-flight (TOF) mass analysers. In MALDI, the sample, co-crystallized with a solid matrix on a target, is desorbed and ionized by laser irradiation. Although designed for the analysis of very high molecular mass compounds, MALDI also plays a role in small molecules analysis, provided that there is no interference from matrix clusters. It has high sensitivity, is easily automated and can give structural information, for example via post-source decay (PSD).[21]

Given the number of analyses required, automatic data interpretation and reporting are necessary to manage the huge amount of data generated. Considerable efforts in this direction are being undertaken.

MS ANALYSIS OF DISCRETE COMPOUNDS IN SOLUTION

In the case of compound arrays in solution, the commonly used analytical techniques are flow-injection or HPLC/MS. Molecular mass information, obtained with a low-cost API, single quadrupole instrument, is the simplest approach to confirm analyte identity. It must be kept in mind, however, that in this case nominal masses are obtained, with accuracy, typically, of only ±0.1 Da. These values cannot be used to assess a molecular formula, for which an accuracy of a few ppm in mass measurements is required.[22] Low-resolution molecular mass information can be considered good structural evidence as long as it is supported by other information, such as preliminary spectroscopic and analytical characterization during the early stage of library development. Several software packages for automatic data handling are commercially available: typically, data systems look for the expected molecular mass in each mass spectrum, and generate a report indicating the presence (or absence) of the desired compound in each vessel, usually in a colour-coded format.

Of course, exact mass measurement of the components of a library adds confidence to their structural assignment: this is especially useful when analysing large libraries, containing many compounds with the same nominal masses. This information can be routinely obtained using TOF or Fourier transform ion cyclotron resonance (FTICR) instruments, equipped with either MALDI or API ion sources. In particular, recent developments in TOF technology, such as reflector TOF analysers, delayed extraction MALDI and API with orthogonal acceleration, have led to significant improvements in resolution (10^4) and mass accuracy (below 10 ppm) with small, relatively low-cost instruments. Two limitations of the measurement are the inability to distinguish isomers and the exponential increase, as the molecular mass increases, of the candidate molecular formulae with masses in the range of the error of measurement. Other aspects relating to exact mass measurements have been extensively discussed.[22,23]

Wang *et al*. reported routine exact mass analysis of combinatorial libraries by HPLC/ESI-TOFMS,[24] with automatic analysis and data processing. To improve accuracy, the authors used a mixture of four mass calibrants, well separated chromatographically, either spiked into each sample or injected after every 10 samples, finding better results with the internal calibration approach. Alternatively, continuous infusion of a mass calibrant during analysis can be performed, via a T-fitting or a double-sprayer ESI source.

Automated sample preparation and MALDI-TOF exact mass analysis of libraries have also been described,[25] making use of angiotensin II and tryptophan–tryptophan as internal calibrants. Matrix, sample and calibrant were automatically spotted on the MALDI target using a patented oscillating capillary nebulizer: more than 80% of measured accuracies were within 5 ppm. High-throughput exact mass determination by ESI-FTICR-MS has also been reported[17] for combinatorial pyrrolamide libraries a mass accuracy below 1 ppm was obtained in the range 100–1400 Da, with a throughput of <30 s per sample and a sample consumption of <1 pmol.

Since a very large number of compounds have often to be characterized, analysis speed is the major concern. In this sense, flow-injection analysis is advantageous compared with HPLC. The introduction of multi-probe autosamplers has led to a dramatic improvement in sample throughput: eight samples are simultaneously injected into eight different

Copyright © 2001 John Wiley & Sons, Ltd.

valves, which are sequentially switched to the inject position, generating eight well-resolved flow-injection peak profiles in about 1 min.[26] With this method, an entire 96-well plate can be analysed in 12 min. Further improvements were reported by Morand *et al*.[27] Starting from slight modifications of commercial devices, they were able to reduce to 5 min the analysis time for a plate, maintaining sample carryover at <1%. The major disadvantage of flow-injection methods is the possibility of in-source discrimination effects due to the co-elution of products with residual reagents, sometimes of much higher ionization yields, and salts. This may lead to undesired phenomena, from clustering to the complete suppression of the analyte signal. A second disadvantage is that it is not always possible to obtain reliable information about sample purity.

These drawbacks are greatly reduced when introducing samples by HPLC, but at the expense of longer analysis times. However, the introduction of high-speed chromatographic columns, based on reduced length (3 cm) and particle size (even <3 μm), has shortened run times to a few minutes or even less, while still maintaining acceptable separation. Generally, these columns are operated at high flow-rates with rapid gradient elution. As an example, Lee *et al*. reported separations of mixtures of eight compounds in 1 min, using a 30 × 4.6 mm i.d. column at a flow-rate of 4 ml min^{-1} and an acetonitrile gradient from 15 to 95% in 0.7 min.[28] Even shorter analysis times were described by Zhang *et al*.,[29] who reported accurate mass measurements and quantitation of five drugs in 18 s with isocratic elution. An important factor to consider is that very narrow chromatographic peaks are generated with these methods, and therefore fast MS scans are needed to ensure adequate chromatographic peak sampling. TOF mass analysers are the instruments of choice, in these cases, owing to their much faster scanning speed compared with quadrupoles.

A further advantage of HPLC is the possibility of working simultaneously with one or more types of detector, connected in series or in parallel with the mass spectrometer. Examples include UV, an evaporative light scattering detector (ELSD)[30] or a chemiluminescent nitrogen detector (CLND).[31] In this way, much better product characterization is achieved, with the addition of complementary information such as sample purity, UV spectra, etc. The use of the recently introduced CLNDs, which is sensitive to the absolute nitrogen content of the sample, looks particularly attractive, giving for each sample the exact relative amount of by-products and the absolute yield. For example, Lewis and Phelps[32] described a combined HPLC-UV/MS-CLND system, providing in a single chromatographic run identity confirmation, purity determination and reaction yields.

In addition to fast HPLC, another promising technique for throughput improvement is parallel LC/MS. This is possible by the use of multiple probe autosamplers and commercially available devices for spraying the mobile phases from four or eight columns into a single API mass spectrometer (MUX technology™).[33] A scheme of the four-channel system is illustrated in Fig. 1: four ESI probes spray their effluents into a sampling rotor, which connects one channel at a time with the entrance aperture of the instrument at appropriate time intervals. The data system is capable of separately acquiring all four channels, giving four distinct data sets. Since the sampling time for each channel is of the order of 100 ms, a critical factor is MS scan speed. The use TOF analysers may be advantageous.

When a library is so large that the analysis of all components is not feasible in a reasonable time, a random

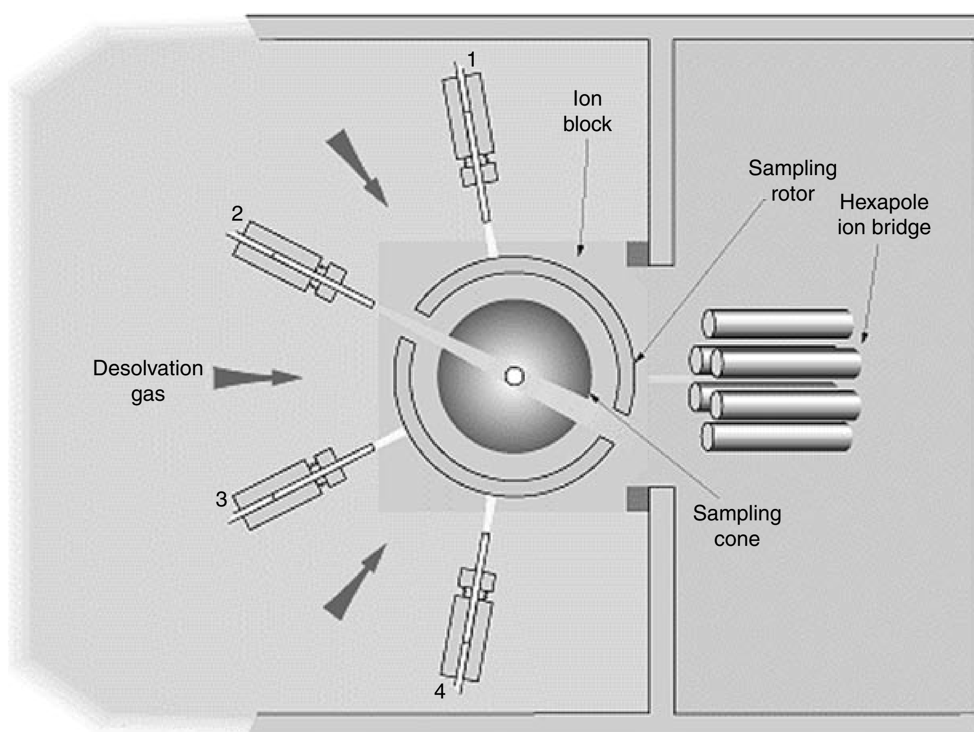

Figure 1. Schematic diagram of the four-channel MUX™ inlet system. Reproduced by permission of www.micromass.co.uk.

Copyright © 2001 John Wiley & Sons, Ltd.

examination of the library is performed (5–7% of the total), using pre-defined sampling protocols.[4] In combinatorial chemistry, analytical strategies are also included in general screening/analysis strategies, according to the individual philosophies of the groups working in this area. Some possible cases are the following:

- total screening followed by analysis of the actives;
- analysis of a subset, total screening, identity and purity confirmation of the actives; and
- characterization and purification before screening.

MASS-DIRECTED PURIFICATION OF LIBRARIES

MS in association with preparative HPLC provides a valid tool for the high-throughput purification of compounds in libraries. In this technique, introduced by Zeng *et al.*,[34] the effluent from a preparative column is split into two parts: 50–100 μL min^{-1} are introduced into an API single quadrupole mass spectrometer and the remaining part is sent to a fraction collector. In normal operation, the effluent is sent to waste, but when the mass spectrometer detects a signal due to the desired product above a pre-set intensity threshold, it triggers its collection until the signal again falls below the threshold. Normally, one fraction per compound is collected in a totally unattended operation. The apparatus is integrated into a unique system, which analyses samples with an analytical column, automatically determines sample purity, then switches to preparative operation and processes the samples having purity below the desired level. Several improvements were introduced to the original system, including the development of a fully automated apparatus for parallel analysis–purification, capable of purifying more than 200 compounds per night.[35] Analytical/preparative systems based on this principle are commercially available. A simplified scheme is shown in Fig. 2.

MS ANALYSIS OF SOLID POLYMER SUPPORTED LIBRARIES

The majority of combinatorial libraries are currently prepared by solid-phase synthesis, commonly carried out on resin beads. The development of techniques capable of analysing organic compounds directly on the solid support is attractive. In principle it should allow the rapid and effective monitoring of reactions, control over the success of the synthesis and structural identification of the bead-bound active compounds. Besides the already-cited case of FTIR, gel-phase and magic angle spinning NMR, which permit direct analysis on the bead, ESI- and MALDI-MS are successfully employed. Egner and Bradley have provided an overview of analytical techniques used in conjunction with solid-phase synthesis.[36] In the case of MS, the first reports described peptide analysis by MALDI following cleavage, usually with TFA vapour, and matrix addition.[37] An interesting simplified procedure makes use of photosensitive linkers to carry out the synthesis. One bead is placed on the MALDI target and matrix is added: simultaneous cleavage and ionization occur under laser irradiation.[13,38] Good results were reported for peptides and oligosaccharides covalently bound to the resin through a photolabile α-methylphenacyl ester linker. However, FMOC-protected peptides were not detected by MALDI analysis, and required previous deprotection.[13] ESI detection of peptides detached from the resin has also been reported.[39,40]

The identification of the bead-bound active compounds is a key objective, for which direct or indirect methods can be used. In both cases, the most important requirements are sensitivity and specificity. Given its sensitivity, MS is currently the method of choice for the direct structural determination of bead-bound compounds, which are typically present in too small amounts to be analysed with other techniques for structural elucidation such as NMR. Successful on-the-bead analysis of small molecules by MALDI has been described.[41] Brummel *et al.* gave a comparative evaluation of ESI, MALDI and TOF-secondary ion MS (SIMS) for the analysis of a bead-bound non-peptidic angiotensin II antagonist, finding

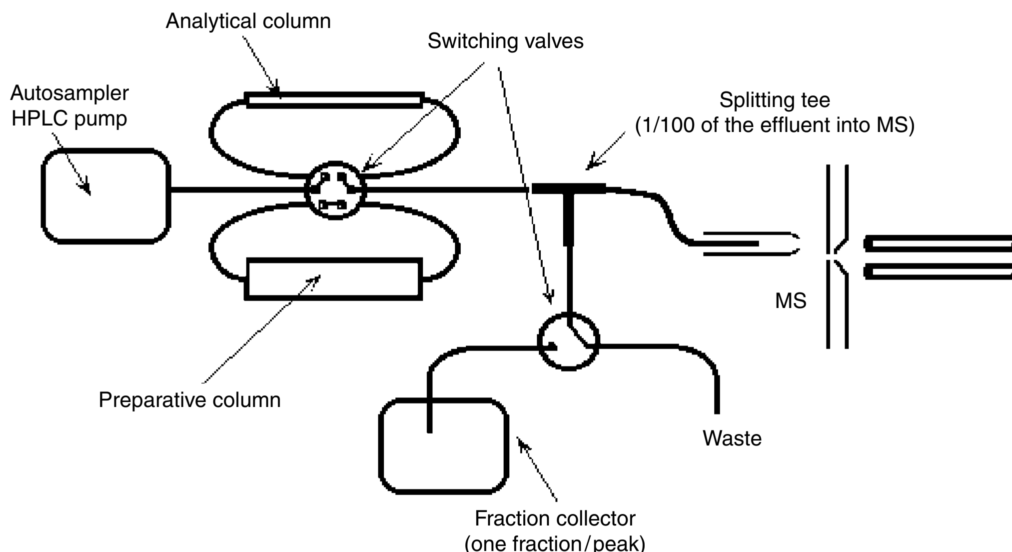

Figure 2. Simplified scheme of a mass-directed analytical–preparative HPLC system for library purification.

Copyright © 2001 John Wiley & Sons, Ltd.

similar results for ESI and MALDI, whereas TOF-SIMS provided the possibility of spatial resolution of the composition of a single bead in the presence of beads carrying different compounds.[42]

Another important issue is specificity: since the number of compounds in a library can be very high, unambiguous discrimination of one out of many candidate structures is sometimes a particularly challenging task. Blom reported on the data precision requirements and the best MS strategies to apply.[43] Interestingly, strategies combining two mass spectral techniques with moderate data accuracy were found to be more specific and robust than high-accuracy measurements performed with a single technique. These principles were applied to dipeptide libraries (1056 components), in which compounds from active beads were photocleaved, dissolved and analysed by LC/ESI-MS, obtaining exact mass and data-dependent product ion spectra in one run. Depending on the mass measurement accuracy required, a hybrid quadrupole-TOF or a triple quadrupole instrument were used. A computer program simulated MS/MS data for all the components, and matched simulated with measured data. This automatic procedure was successful in 80% of the analyses.[44]

Another elegant approach, but one limited to peptide libraries, is to add a small amount of a terminating reagent at each step of the synthesis: after cleavage, a mixture containing the desired peptide plus small amounts of the peptide fragments is obtained, and the peptide sequence is easily determined by ladder sequencing via MALDI-TOF analysis.[45]

MS is also applied in some indirect methods of identification of the actives. Indirect methods are particularly useful when faced with very large libraries (in which direct methods fail to give unique results because of high molecular mass degeneracy) or when there are detection problems for the compounds under study. In this case, a second molecule is linked to each bead in addition to the library component, using orthogonal chemistry. This molecular tag maintains in coded form the full information of the compound structure. It is released from the bead after biological testing and determined with high sensitivity using a convenient analytical method, not necessarily MS-based. Some methods used are GC with electron-capture detection[46] for the detection of volatile halocarbon tags released from the beads via oxidative cleavage, or reversed-phase HPLC for dansylated secondary amine tags[47] released by mineral acid hydrolysis. In other instances, non-chemical codes are used, as in a radiofrequency-based system.[48] MS-based decoding strategies making use of tags labelled with ^{13}C or deuterium have been described.[49–52]

Of course, indirect methods do not ensure that the corresponding compounds are really present on the bead, nor do they give any indication about compound purity. A method to circumvent this potential problem is the analytical construct technology.[49] In this method, the tag and the compound are incorporated with two orthogonal linkers in the same construct having the structure linker 1–tag–linker 2–compound. Linker 1 connects the construct to the resin, and the tag is encoded with stable isotopes. It may also contain an MS sensitizer to allow easy detection by ESI or MALDI. When linker 2 is cleaved, the compound is released for biological testing. Linker 1 is cleaved for MS analysis of the entire construct for quality control or decoding purposes.[51]

MS ANALYSIS OF POOLED LIBRARIES

The synthesis of pooled libraries often generates very complex mixtures, containing the compounds of interest, as well as eventual by-products, residual reagents and impurities. In assessing identity and purity of a pooled library, an analytical method should give answers to the following questions:

- Are all the components present?
- Are the structures consistent with those expected?
- Are there any synthesis by-products, residual reagents or other extraneous substances? What are their structures?
- Are the components present in equimolar amounts? What is their absolute amount?

The first obvious approach for the analytical characterization of such libraries is trying to confirm the presence of all the expected components. The potential of MS methods for analysis of synthetic peptide libraries was investigated for the first time by Metzger *et al.*, analysing an equimolar mixture of 48 octapeptides by ESI-MS.[53] It was shown that tandem MS techniques allowed rapid detection and characterization of products and also by-products arising from incomplete deprotection. The potential and limits of MS-based techniques for analysing larger peptide libraries were discussed in more detail by the same group.[54] A similar study, addressed to small-molecule libraries, was reported by Dunayevskiy *et al.*:[55] starting from the point of view that the characterization of mixtures containing many thousand compounds was unrealistic, a detailed analysis was performed on model libraries of up to 55 components by combining positive/negative ion ESI and different MS/MS techniques with a triple quadrupole instrument.

In the case of even slightly larger pooled libraries, a frequent problem is to discriminate between nominally isobaric components. It is evident that high mass resolution is crucial in this case. For example, the mass separation of two compounds having a nominal molecular mass of 500 Da and differing by a lysine at the place of a glutamine residue would require a resolution of about 14 000. Such a critical requirement reduces the potentially useful mass analysis devices almost exclusively to FTICR instruments.

Fang *et al.*[56] described ESI-FTICR analysis of libraries of 36, 78 and 120 components: 70–80% of the expected structures were confirmed by determining their exact molecular masses without chromatographic separation. A limitation indicated by the authors was the possible overloading of the ICR cell when directly infusing a relatively large library, resulting in loss of resolution and accuracy due to space charge effects. Coupling of HPLC or capillary zone electrophoresis (CZE) with ESI-FTICR was suggested as a possible solution to this problem. The advantages of LC/MS were recently demonstrated by Schmid *et al.*,[57] who followed both a direct infusion and a micro-LC approach for the

Copyright © 2001 John Wiley & Sons, Ltd.

ESI-FTICR analysis of a 144-member pyrazole library. They achieved in this case complete assignment, whereas only 93% of the components were detected by direct infusion. The HPLC separation allowed the detection of isomers, and avoided some ion suppression effects in the ion source.

A second general method for characterizing pooled libraries does not rely on the detection of every single component, but considers the library as a whole, and uses mostly low-resolution instruments. The simplest approach is analysis by ESI-MS with a direct infusion or flow-injection method, obtaining the mass spectrum of the library. The mass distribution of the spectrum is compared with the theoretical distribution, which can be calculated using computer programs.[58,59] Of course, this investigation cannot detect every single component, since in low-resolution instruments all the compounds with the same nominal mass appear as a single peak; however, it provides instead a fingerprint spectrum giving qualitative information. The comparison is based on some assumptions, i.e. that all the library components give the same molar response to mass detection, the molecular ion abundances of isobaric components are additive and there is no signal suppression of individual compounds. These conditions are far from being satisfied in actual ESI mass spectra, in which, besides the molecular ions, the compounds of interest are represented by adduct ions, isotopic ions, fragment ions, multiply charged ions and cluster ions. In addition, their molar responses may vary greatly even with small structural differences. Despite this, such an approach can provide a quick and effective evaluation of combinatorial mixtures. A recent example was reported by Yates *et al.*:[60] using flow-injection ESI with a single quadrupole mass spectrometer, they characterized with high-throughput libraries of up to 400 components by matching the measured with theoretical spectra through automatic correlation analysis. For each sample, such an analysis gave a compound score to confirm library products, and a purity score to indicate the occurrence of by-products. The authors report that ion suppression was not a significant problem for the libraries under study. This was not the case, instead, for a series of pooled libraries containing $n \times 324$ components ($n = 1, 2$), analysed in our laboratory. Preliminary studies performed by positive and negative ion ESI with direct infusion showed significant differences in ionization yields, resulting, not surprisingly, in higher sensitivities for the more basic compounds in positive ESI and for the more acidic compounds in negative ESI. To solve this problem partly, we chose HPLC/MS as the first approach. A second advantage would have been that, since the libraries contained many possible isobaric species (up to 10), on-line HPLC would have ensured at least a rough separation of isobaric substances. Given that we also wanted to acquire product ion MS/MS data to confirm structures and assign unknowns, we used an ion trap mass analyser, more sensitive than a triple quadrupole in full-scan MS/MS, and capable of switching from MS to MS/MS automatically selecting the parent ion. As a third advantage, an HPLC separation would have maximized the amount of structural information obtained in this mode of operation.

Fig. 3(A) shows the contour plot corresponding to the HPLC/MS analysis of a typical library in which the component masses range from 388 to 600 Da. Ideally, each spot corresponds to one compound, allowing resolution by molecular mass of many co-eluting substances. Contour plots of libraries give at a glance a lot of information about their quality, as discussed by Metzger *et al.*;[54] however, since the count of the spots is impractical, it is difficult to assess from the map that all the components are represented. Extracted

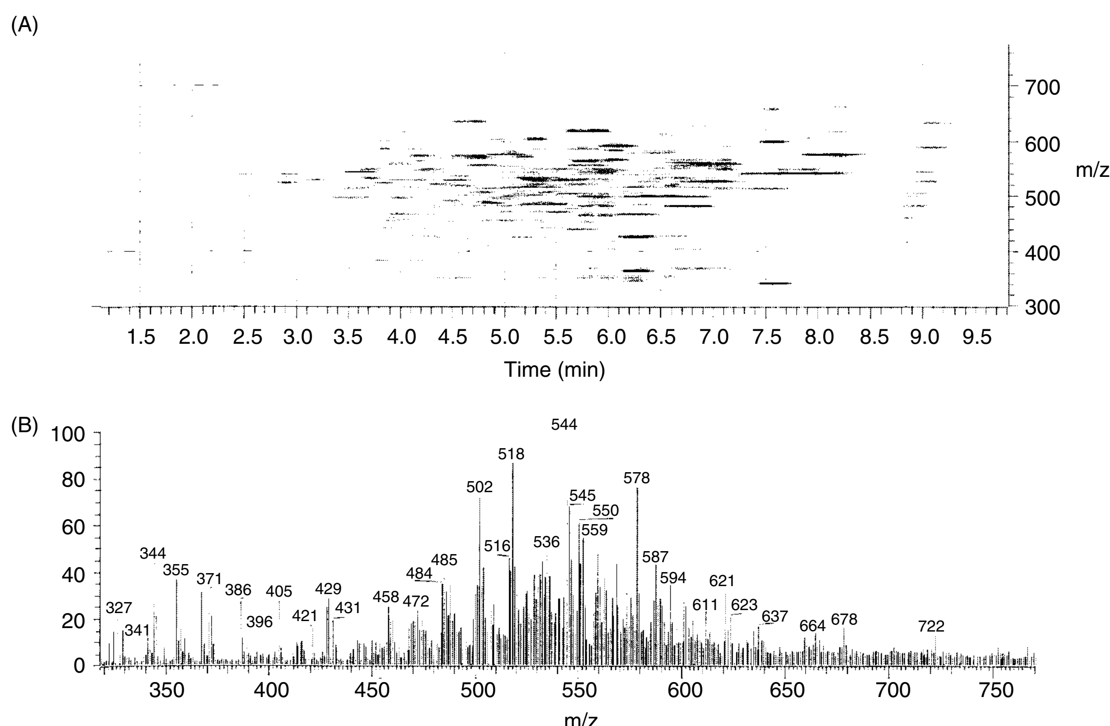

Figure 3. Contour plot (A) and average mass spectrum (B) of a typical library.

Copyright © 2001 John Wiley & Sons, Ltd.

JMS

chromatograms of some randomly chosen expected ions may be useful, provided that all the isobaric components are chromatographically resolved, which may not always be the case.

The simulated spectrum of the library, obtained with a spreadsheet program, is compared with the average mass spectrum of the library in Fig. 3(B). Very similar patterns are shown, giving good confidence about the success of the synthesis. However, the peak relative abundances differ significantly from those expected from the simulated spectrum, in which all the components are assumed to have equal molar response, since, as stated before, this condition is not always met in an ESI ion source. Some extraneous peaks are observed in the spectrum; those in the lower mass portion are assignable to fragment ions of the library, whereas those at higher masses belong, in this case, to traces of ubiquitous polyethylene glycol-based contaminants.

With the principal aim of ensuring the occurrence of all the desired components, a second method was then developed: it employed a triple quadrupole instrument operated either in the precursor or in the constant neutral losses scan mode. As already shown,[55] by proper choice of the product ion or the neutral loss in the expected MS/MS fragmentation pattern of the components, it is possible to detect in the whole library only the compounds with the selected structural characteristics.

An example is a library of heterogeneous oligomers of general formula X-aa1-Y-aa2-Z, where aa1 and aa2 are one of 18 amino acids, whereas X, Y and Z are fixed. Figure 4(A) and (B) show, respectively, the mass spectrum of the whole library and the precursor ion spectrum of *m/z* 455, a fragment expected in the product ion tandem mass spectra of all the compounds having Arg as the aa1 and different aa2. Precursor ion scanning selects the components whose fragmentation generated this ion, showing that all the aa2 amino acids are represented (considering that the couples Lys–Gln and Leu–Ile are isobaric and cannot be mass-resolved with the instrument used, and that the signal of the compound with two Args is very low). In a similar MS/MS experiment, the occurrence of all the aa1 amino acids for a given aa2 was demonstrated.

This method cannot give quantitative information, and its applicability is not general, depending on the type of compound and its fragmentation; nevertheless, it provides a very rapid evaluation of the composition also in complicated mixtures, and gives complementary information with respect to the LC/MS method described above. The latter, however, is more generally applicable and is therefore more amenable to routine use. Furthermore, it gives automatically MS/MS data in the same chromatographic run, as stated before. This is a very rapid and effective method to acquire structural information on selected compounds, and may be applied to confirm expected structures and also to elucidate unknown, undesired products.

As an example, Fig. 5(A) and (B) show, respectively, the LC/MS contour plot and the average spectrum of another library of the same general formula X-aa1-Y-aa2-Z as above, whose expected molecular masses range from 356 to 568 Da.

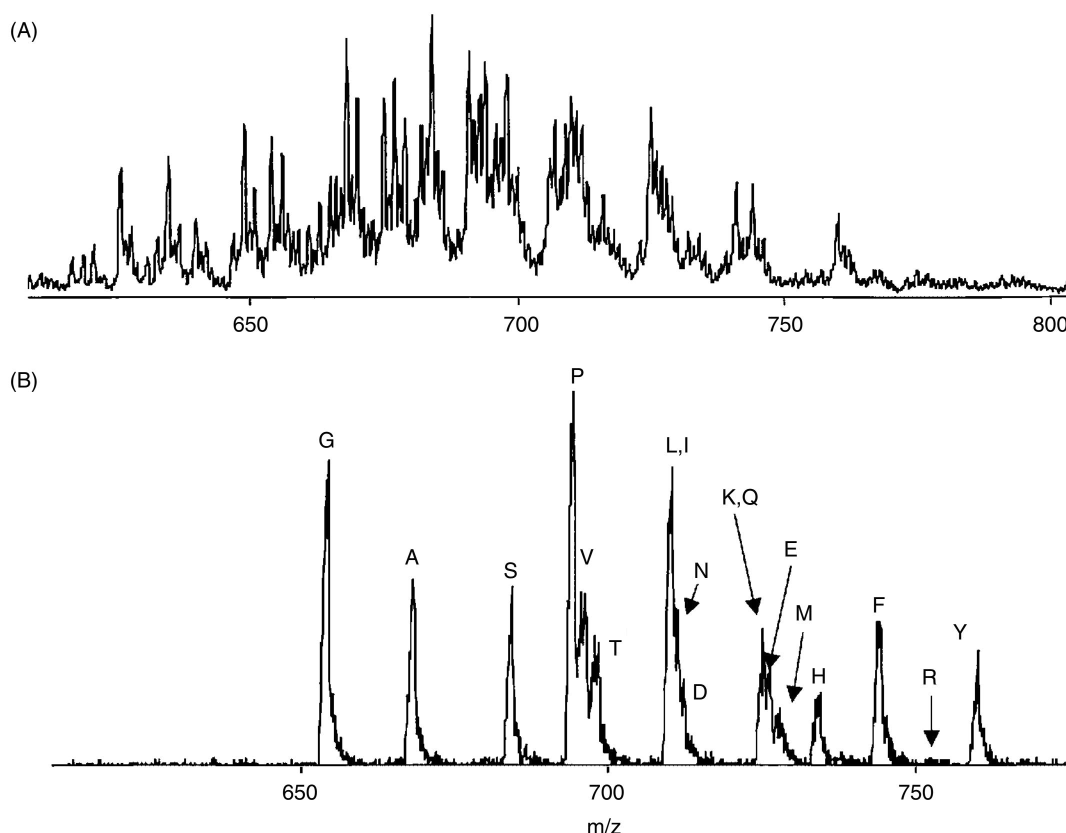

Figure 4. Mass spectrum (A) and precursor ion tandem mass spectrum of *m/z* 455 (B) of a library. The letters correspond to the single letter codes of aa2 amino acids (see text for details).

Copyright © 2001 John Wiley & Sons, Ltd.

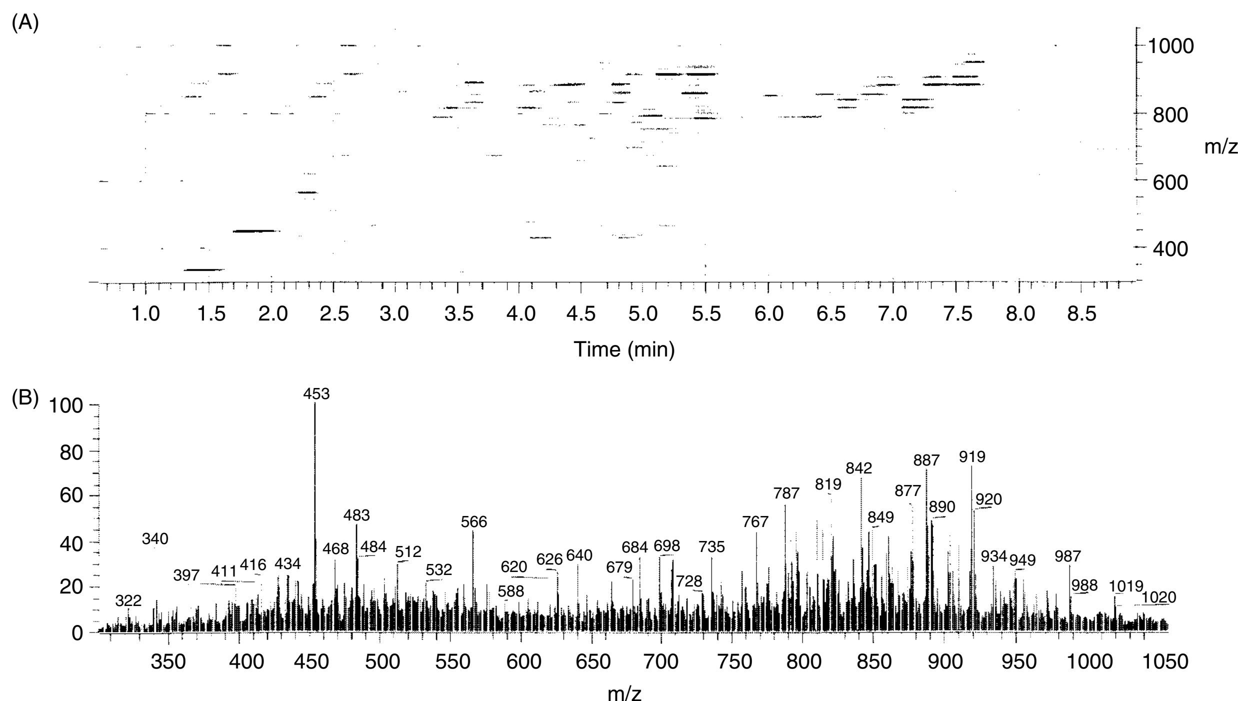

Figure 5. Contour plot (A) and average mass spectrum (B) of an unsuccessfully synthesized library.

In this case, X was a succinyl group. In addition to the expected library, which was also confirmed by MS/MS analysis of some expected molecular ions, there was a second distribution of peaks centred on m/z 887, having retention times similar to those of the desired peaks. The analysis of the product ion tandem mass spectra of some extraneous peaks revealed that all the additional compounds had the same general formula Z-aa2-Y-aa1-CO-$(CH_2)_2$-CO-aa1-Y-aa2-Z, in which a succinyl bridged two oligomeric chains. This information was returned to the synthetic chemists, allowing them to design a new synthetic method in order to avoid the formation of this undesired 'side-library'.

MS-BASED SCREENING OF LIBRARIES IN SOLUTION

Identification of the active compounds in mixture libraries is normally performed by deconvolution strategies, in which iterative synthesis and assay of sub-libraries are performed.[61] Since this is a labour-intensive and time-consuming procedure, alternative deconvolution methods have been developed, some of which involve MS. An overview of deconvolution approaches has been given by Schriemer and Hindsgaul.[62] Essentially, the majority of the alternative assays are based on mixture incubation with the target, separation of bound from non-bound ligands, eventual disruption of the target–ligand association and ligand detection/identification with a suitable method.

Size-exclusion chromatography (SEC) is one of the techniques used for the separation of ligand–receptor complexes from non-bound library components. In an off-line approach, Kaur *et al.*[63] incubated the target protein (500 pM) with libraries of up to 600 peptoids (500 pM each). Following SEC and desalting/complex dissociation using a peptide cartridge, the ligands were eluted into a triple quadrupole mass spectrometer with ESI. An automated multi-dimensional chromatographic system based on SEC and reversed-phase separation coupled to ESI was presented by Hsieh *et al.*[64]

Another technique involves immunoaffinity ultrafiltration combined with ESI-MS.[65,66] Antibodies raised against selected benzodiazepines were used to isolate them from a library of about 20 compounds. Antigen–antibody complexes were separated by centrifugal ultrafiltration, dissociated by acidification and analyzed by LC/MS.

Based on the same separation technique is pulsed ultrafiltration:[67] incubation of a library with the receptor is performed, mostly on-line, in an ultrafiltration chamber. After elution of non-binding compounds, ligands are released by denaturation and detected by MS or reversed-phase HPLC/MS. Application examples were reported for dihydrofolate reductase[68] and adenosine deaminase inhibitors.[67] Other reported methods based on on-line separation of bound ligands from low-affinity molecules are affinity capillary electrophoresis[69] and frontal affinity chromatography.[70]

Gao *et al.* described a technique, entirely based on MS, called bioaffinity characterization MS.[71] In this procedure, the entire library–target mixture formed after incubation is injected into an ESI-FTICR mass spectrometer, without any separation. The non-covalent target–ligand complex is isolated in the ICR cell, dissociated by collisional activation and the ligands are analysed by a further step of mass selection and collisional activation. The method was used to screen two libraries (289 and 256 compounds) of carbonic anhydrase II inhibitors. Good correlation was found between the relative ion abundances of ligands and their affinity constants in solution.

Copyright © 2001 John Wiley & Sons, Ltd.

All the above methods rely on non-covalent interactions of molecules with a target; an MS method for measuring covalent complexes was described, screening the antiaggregatory activity of isothiocyanates for sickle haemoglobin. The applicability of the above-described methods, in which affinity selection is coupled with MS, is limited to a few cases owing to a series of drawbacks.[1] Essentially, they require larger amounts of target protein than biochemical assays, and this target must be highly purified and soluble. They are attractive, however, because they allow direct structural identification of the active components, without requiring lengthy deconvolution procedures.

CONCLUSION

MS-based techniques are extensively employed in all the aspects of combinatorial technologies aimed at the discovery of new drugs. This is due to some intrinsic characteristics of MS, i.e. sensitivity and specificity, as well as to recent technological progress, leading to affordable and easy-to-use instruments, which meet the needs of high-throughput, sensitive and informative analyses underlying all the steps in modern drug discovery. Successful application of MS in combinatorial chemistry, however, does not depend uniquely on its favourable characteristics: even more, it relies on effective interaction and integration with other disciplines, such as separation sciences, robotics and information management. We will probably assist in the next few years to increase efforts in this direction.

Acknowledgements
We thank Dr Andrea Raffaelli (Consiglio Nazionale delle Ricerche) for triple quadrupole mass measurements and Dr Lisa Taverni for assistance with bibliographic documentation.

REFERENCES

1. Floyd CD, Leblanc C, Whittaker M. *Prog. Med. Chem.* 1999; **36**: 91.
2. Tiebes D. In *Combinatorial Chemistry*, Jung G (ed.). Wiley-VCH: Weinheim, 1999; 1.
3. Fassina G, Miertus S. In *Combinatorial Chemistry and Technology*, Miertus S, Fassina G (eds). Marcel Dekker: New York, 1999; 1.
4. Furka A, Sebestyen M, Asgedom M, Dibo G. *Int. J. Pept. Protein Res.* 1991; **37**: 487.
5. Sepetov N, Issakova O. In *Combinatorial Chemistry and Technology*, Miertus S, Fassina G (eds). Marcel Dekker: New York, 1999; 169.
6. Egner BJ, Bradley M. *Drug Discov. Today* 1997; **2**: 102.
7. Kassel DB. *Chem. Rev.* 2001; **101**: 255.
8. Look GC, Holmes CP, Chinn JP, Gallop MA. *J. Org. Chem.* 1994; **59**: 7588.
9. Stover HDH, Frechet JMJ. *Macromolecules* 1989; **22**: 1574.
10. Warrass R, Lippens G. In *Combinatorial Chemistry*, Jung G (ed.). Wiley-VCH: Weinheim, 1999; 533.
11. Chin J, Fell JB, Jarosinski M, Shapiro MJ, Wareing JR. *J. Org. Chem.* 1998; **63**: 386.
12. Yan B, Gremlich HU, Moss S, Coppola GM, Sun Q, Liu L. *J. Comb. Chem.* 1999; **1**: 46.
13. Siuzdak G, Lewis JK. *Biotechnol. Bioeng.* 1998; **61**: 127.
14. Süssmuth R, Trautwein A, Richter H, Nicholson G, Jung G. In *Combinatorial Chemistry*, Jung G (ed.). Wiley-VCH: Weinheim, 1999; 500.
15. Süssmuth R, Jung G. *J. Chromatogr. B* 1999; **725**: 49.
16. Enjalbal C, Martinez J, Aubagnac GL. *Mass Spectrom. Rev.* 2000; **19**: 139.
17. Walk TB, Trautwein A, Jack R, Thyroff M, Jertz R, Baykut G, Jung G, In *Proceedings of the 47th ASMS Conference on Mass Spectrometry and Allied Topics*, Dallas TX, 1999.
18. Fenn JB, Mann M, Meng CC, Wong SF. *Mass Spectrom. Rev.* 1990; **9**: 37.
19. Duffin KL, Wachs T, Henion JD. *Anal. Chem.* 1992; **64**: 61.
20. Karas M, Hillenkamp F. *Anal. Chem.* 1988; **60**: 2299.
21. Vestal ML, Juhasz P, Martin SA. *Rapid Commun. Mass Spectrom.* 1995; **9**: 1044.
22. Biemann K. *Methods Enzymol.* 1990; **193**: 295.
23. Russell DH, Edmondson RD. *J. Mass Spectrom.* 1997; **32**: 263.
24. Wang X, Takach E, Deutschman R, Zhou J, Kassel DB. In *Proceedings of the 47th ASMS Conference on Mass Spectrometry and Allied Topics*, Dallas TX, 1999.
25. Lake DA, Johnston MV, Larsen BS, McEwen CN. In *Proceedings of the 47th ASMS Conference on Mass Spectrometry and Allied Topics*, Dallas TX, 1999.
26. Wang T, Zeng L, Strader T, Burton L, Kassel DB. *Rapid Commun. Mass Spectrom.* 1998; **12**: 1123.
27. Morand KL, Burt TM, Regg BT. In *Proceedings of the 48th ASMS Conference on Mass Spectrometry and Allied Topics*, Long Beach, CA, 2000.
28. Lee H, Li L, Kyranos J. In *Proceedings of the 47th ASMS Conference on Mass Spectrometry and Allied Topics*, Dallas TX, 1999.
29. Zhang H, Heinig K, Henion J. *J. Mass Spectrom.* 2000; **35**: 423.
30. Kibbey CE. *Mol. Diversity* 1995; **1**: 247.
31. Bizanek R, Manes JD, Fujinari E. *Pept. Res.* 1996; **9**: 40.
32. Lewis K, Phelps D. In *Proceedings of the 47th ASMS Conference on Mass Spectrometry and Allied Topics*, Dallas TX, 1999.
33. Sage AB, Bowers G, Chandrasurin P, Castro-Perez J, Preece S. In *Proceedings 48th ASMS Conference on Mass Spectrometry and Allied Topics*, Long Beach CA, 2000.
34. Zeng L, Burton L, Yumg K, Shushan B, Kassel BD. *J. Chromatogr. A* 1998; **749**: 3.
35. Zeng L, Kassel DB. *Anal. Chem.* 1998; **70**: 4380.
36. Egner BJ, Bradley M. *Drug Discov. Today* 1997; **2**: 12.
37. Egner BJ, Langley GJ, Bradley M. *J. Org. Chem.* 1995; **60**: 2652.
38. Fitzgerald MC, Harris K, Shevlin CG, Siuzdak G. *Bioorg. Med. Chem. Lett.* 1996; **6**: 979.
39. Brown BB, Wagner DS, Geysen HM. *Mol. Diversity* 1995; **1**: 4.
40. McKeown SC, Watson SP, Carr RAE, Marshall P. *Tetrahedron Lett.* 1999; **40**: 2407.
41. Haskins NJ, Hunter DJ, Organ DJ, Rabman SS, Thom C. *Rapid Commun. Mass Spectrom.* 1995; **9**: 1437.
42. Brummel CL, Vickerman JC, Carr SA, Hemling ME, Roberts GD, Johnson W, Weinstock J, Gaitanopoulos D, Benkovic SJ, Winograd N. *Anal. Chem.* 1996; **68**: 237.
43. Blom KF. *Anal. Chem.* 1997; **69**: 4354.
44. Blom KF, Combs AP, Rockwell AL, Oldemburg KR, Zhang JH, Chen T. *Rapid Commun. Mass Spectrom.* 1998; **12**: 1192.
45. Youngquist RS, Fuentes GR, Lacey MP, Keough T. *J. Am. Chem. Soc.* 1995; **114**: 3900.
46. Nestler HP, Bartlett PA, Still WC. *J. Org. Chem.* 1994; **59**: 4723.
47. Ni ZJ, Maclean D, Holmes CP, Murphy MM, Ruhland B, Jacobs JW, Grovdon EM, Gallop MA. *J. Med. Chem.* 1996; **39**: 1601.
48. Li W, Czarnik AW, Xiao XY, Lillig J. *J. Comb. Chem.* 2000; **2**: 224.
49. Geysen HM, Wagner CD, Bodnar WM, Markworth CJ, Parke GJ, Schoenen FJ, Wagner DS, Kinder DS. *Chem. Biol.* 1996; **3**: 679.
50. Wagner DS, Markworth CJ, Wagner CD, Schoenen FJ, Rewerts CE, Kay BK, Geysen HM. *Comb. Chem. High Throughput Screening* 1998; **1**: 143.
51. Lorthioir O, Carr RAE, Congreve MS, Geysen MH, Kay C, Marshall P, McKeown SC, Parr NJ, Scicinski JJ, Watson SP. *Anal. Chem.* 2001; **73**: 963.
52. Lane SJ, Pipe A. *Rapid Commun. Mass Spectrom.* 2000; 14.

Copyright © 2001 John Wiley & Sons, Ltd.

53. Metzger JW, Kempter C, Wiesmüller KH, Jung G. *Anal. Biochem.* 1994; **219**: 261.
54. Metzger JW, Wiesmuller KH, Kienle S, Brunjes J, Jung G. In *Combinatorial Peptide and Non Peptide Libraries—a Handbook*, Jung G (ed). VCH: Weinheim, 1996; 247.
55. Dunayevskiy Y, Vouros P, Carell T, Wintner EA, Rebek J. *Anal. Chem* 1995; **67**: 2906.
56. Fang AS, Vouros P, Stacey CC, Kruppa GH, Laukien FH, Wintner EA, Carell T, Rebek J. *Comb. Chem. High Throughput Screening* 1998; **1**: 23.
57. Schmid DG, Grosche P, Jung G. *Rapid Commun. Mass Spectrom.* 2001; **15**: 341.
58. Brünjes J, Metzger JW, Jung G. In *Combinatorial Peptide and Non Peptide Libraries—a Handbook*, Jung G (ed.). VCH: Weinheim, 1996; 511.
59. Demirev PA, Zubarev RA. *Anal. Chem.* 1997; **69**: 2893.
60. Yates N, Wislocki D, Roberts A, Berk S, Klatt T, Shen DM, Willoughby C, Rosauer K, Chapman K, Griffin P. *Anal. Chem.* 2001; **73**: 2941.
61. Konings DAM, Wyatt JR, Ecker DJ, Freier SM. *J. Med. Chem.* 1997; **40**: 4386.
62. Schriemer DC, Hindsgaul O. *Comb. Chem. High Throughput Screening* 1998; **1**: 155.
63. Kaur S, McGuire L, Tang D, Dollinger G, Huebner V. *J. Protein Chem.* 1997; **16**: 505.
64. Hsieh YF, Gordon N, Regnier F, Afeyan N, Martin SA, Vella GJ. *Mol. Diversity* 1996; **2**: 189.
65. Wieboldt R, Zweigenbaum J, Henion J. *Anal. Chem.* 1997; **69**: 1683.
66. Wieboldt R, Zweigenbaum J, Henion JD. In *Mass Spectrometry of Biological Materials*, Larsen BS, McEwen CN (eds). Marcel Dekker: New York, 1998; 81.
67. Van Breemen RB, Huang CR, Nikolic D, Woodbury CP, Zhao YZ, Venton DL. *Anal. Chem.* 1997; **69**: 2159.
68. Nikolic D, Van Breemen RB. *Comb. Chem. High Throughput Screening* 1998; **1**: 47.
69. Chu YH, Dunayevskiy YM, Kirby DP, Vouros P, Karger BL. *J. Am. Chem. Soc.* 1996; **118**: 7827.
70. Schriemer DC, Bundle DR, Li L, Hindsgaul O. *Angew. Chem., Int. Ed. Engl.* 1998; **37**: 3383.
71. Gao J, Cheng X, Chen R, Sigal GB, Bruce JE, Schwartz BL, Hofstadler SA, Anderson GA, Smith RD, Whitesides G. *J. Med. Chem.* 1996; **39**: 1949.
72. Park S, Wanna L, Johnson ME, Venton DL. *J. Comb. Chem.* 2000; **2**: 314.

Copyright © 2001 John Wiley & Sons, Ltd.