

(19) World Intellectual Property  
Organization  
International Bureau



(43) International Publication Date  
14 October 2004 (14.10.2004)

PCT

(10) International Publication Number  
**WO 2004/088536 A2**

(51) International Patent Classification<sup>7</sup>: **G06F 17/30**

(21) International Application Number:  
PCT/GB2004/001228

(22) International Filing Date: 23 March 2004 (23.03.2004)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
0307352.5 29 March 2003 (29.03.2003) GB

(71) Applicant (for all designated States except US): **QINETIQ LIMITED** [GB/GB]; Registered Office, 85 Buckingham Gate, London SW1E 6PD (GB).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **HEAL, Richard, D., A.** [GB/GB]; Qinetiq Limited, Winfrith Technology Centre, Winfrith Newburgh, Room A22/140, Dorchester, Dorset DT2 8XJ (GB). **PARSONS, Alan, T.** [GB/GB]; Qinetiq Limited, Winfrith Technology Centre, Winfrith Newburgh, Room A22/42, Dorchester, Dorset DT2 8XJ (GB).

(74) Agent: **TOCHER, A., J.**; QinetiQ Ltd, IP Formalities, Cody Technology Park, A4 Building, Room G016, Ively Road, Farnborough, Hampshire GU14 0LX (GB).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

**Declaration under Rule 4.17:**

— of inventorship (Rule 4.17(iv)) for US only

**Published:**

— without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: IMPROVEMENTS IN AND RELATING TO THE ANALYSIS OF COMPOUNDS

(57) Abstract: The analysis of compounds utilising the physical response of cellular networks is described. Typically, the electrical characteristics of the cellular network are monitored although other characteristics such as fluorescence may be monitored. In any event, the analysis utilises signal processing techniques to derive a set of features which may be evaluated against a library of known responses. The technique may be applied to both the detection and identification of unknown compounds and the detection of concentrations of known compounds.



WO 2004/088536 A2

Improvements in and relating to the analysis of compounds

5 The present invention relates to the analysis of compounds, particularly, although not exclusively chemical compounds utilised in the healthcare, pharmaceutical, cosmetic and environmental sectors.

10 It has become the case that in the search for ever more effective pharmaceutical compounds an ever-increasing amount of time, effort and resources have been devoted to identifying and isolating potentially beneficial chemical compounds. Traditionally, the approach has been to select a molecular target within a biochemical pathway, such as an enzyme or a receptor where interaction with the target by a compound could lead to changes which treat the disease. Typically, the interaction would take the form of the compound inhibiting or exciting the pathway. Clearly, a large number of targets will be  
15 under investigation at any one time. In order to evaluate the target against potentially useful compounds it is necessary to produce samples of the target for testing, typically through a cloning process. The target is then screened in a series of tests against these compounds with a view to eliminating those compounds which are unsuitable and to identify those compounds that are potentially valuable. It is sometimes the case that  
20 there may be sufficient biostructural information on the molecular target to suggest the design of potentially valuable compounds. Even so, for the most part hundreds of thousands of compounds are typically screened using robotic technology. Typically, the entire process from initial selection of a target through to the identification of candidate compounds can take several years.

25

Once identified as potentially of value in this first screening phase, compounds showing the appropriate activity are subjected to further screens with the aim of determining their level of potency and selectivity for the target. From these data, leads will be identified.

30 Once a potential candidate compound has been identified, it is then subjected to further development including more screening to meet the needs of various studies both clinical and non-clinical studies. The biological effects of a compound will be assessed, wherever possible avoiding using animals in safety testing. Thus, cells in culture are an attractive alternative for the basis for such investigations. Increasingly automation is  
35 being applied to such assessment and whilst for the most part conventional assaying techniques are utilised there has been some initial attempts at employing automated techniques.

One such technique which has been applied to the analysis of cell culture in response to a compound is that set out in US Patent No. 6,377,057 which describes a technique and apparatus for classifying biological agents according to the spectral density signature of evoked changes in cellular electric potential. It is suggested in the Patent that the approaches it teaches are intended to go beyond those previous attempts to measure cellular electric potential. Such early attempts have, it is suggested, produced output more suited to interpretation by an experienced neuroscientist. Indeed, although such tools have been available to researchers and expert practitioners such as cardiologists since the early 1970's, it is suggested that the invention disclosed in the Patent is intended to be of more general use. As such the patent discloses a relatively unsophisticated analysis mechanism based on interpreting the power spectral density (PSD) of a cellular response. Thus, whilst the Patent teaches that the technique is capable of identifying unknown compounds and determining the characteristics of compounds, analysis based purely on the spectral density changes of such evoked membrane potential or action potential is considered to limit the value of the results obtained in the interests of reducing the complexity of analysis.

According to a first aspect of the invention, there is provided a compound analysis method: this method comprises determining a vector quantity having a number of dimensions equal to a number of features derived from the electrical output of an electrically active cellular network, with each component of said vector being representative of a change in said feature resulting from the application of a compound to said electrically active cellular network and classifying said vector in accordance with a predetermined cluster analysis methodology.

When stimulated, cellular activity may change. By analysing the changes in the signal and comparing with a library of pre-determined behavioural features, the stimulus can be characterised and/or the impact on the cellular network quantified. Using processing algorithms based on those currently used for sonar signal processing, signal features can be related to molecular events occurring within the cellular structure. Such features may be electrical, chemical or fluorescent in nature, the requirement being that a suitable transducer is provided.

According to another aspect of the invention, there is provided a compound analysis system, the system comprising a micro-electrode array provided by a bio-compatible

substrate having a plurality of electrodes situated thereon, said electrodes having an arrangement on said substrate corresponding substantially to that of an electrically active cellular network disposable in use thereon a multi-channel amplifier coupled to said electrodes and an analyser operatively connected to said amplifier to determine for each  
5 active channel a vector quantity having a number of dimensions equal to a number of features derived from the electrical output of said electrically active cellular network with each component of said vector being representative of a change in said feature.

Conveniently, the system may be used for the testing of blood/ urine samples for drugs.  
10 A particular application may be to test for performance enhancing drugs in athletes or for patient compliance.

The delivery of a compound to the system may be placed under the control of a perfusion system. By utilising such a perfusion system it is possible to be more confident  
15 that accurate delivery of a compound, such as a drug takes place. A perfusion system allows correct and potentially varying doses of the drug, to be delivered at a know time. Furthermore, the perfusion system may ensure that the drug is added at the correct temperature with a controlled and smooth delivery.

20 Thus, the system is particularly suitable for the needs of the pharmaceutical market, in that it may identify and quantify behavioural changes in the cellular structure when exposed to external stimuli. Conveniently, post-processing activities could include comparison with a pre-determined library of responses to identify the stimulator, or a simple output to determine whether the reaction falls within acceptable limits

25 According to a still further aspect of the invention, there is provided a compound analysis apparatus, the apparatus including a processor and a memory, the processor being operable in response to signals derived from a micro-electrode array connected, in use, thereto, to determine a vector quantity having a number of dimensions equal to a number  
30 of features derived from the electrical output of said micro-electrode array with each component of said vector being representative of a change in said feature, wherein said memory contains a library of features characterising known compounds such that classification of said vector enables identification of a compound deposited, in use, on said array, in accordance with a predetermined measure of statistical reliability.

The signals derived from the micro-electrode array although electrical in nature may originate from not only electrical activity but also other changes in the physical characteristics of a cellular network deposited, in use, on the array. Clearly, there is a requirement that a transducer on or in proximity to the array is capable of providing  
5 electrical signal representative of the characteristics being monitored.

According to still another aspect of the invention, there is provided a sensor for compound detection, the sensor comprising a receptacle for a micro-electrode array, said receptacle having a connector for receiving electrical signals from said array when  
10 received in said receptacle, an amplifier for amplifying said signals and a processor, the processor being operable in response to said signals to determine a vector quantity having a number of dimensions equal to a number of features derived from the electrical output of said micro-electrode array with each component of said vector being representative of a change in said feature.

15 Such a sensor may be packaged to suit a variety of applications such as, but not exclusively, water quality analysis, environmental monitoring and process control in, for example, the production of pharmaceuticals, cosmetics, foodstuffs and the like.

20 In general, once a set of baseline measurements from an array has been recorded, any change in cellular behaviour can be identified and quantified. Such changes may be induced through exposure to a stimulus. In the embodiments of the invention set out below, the stimulus would be in the form of a chemical compound.

25 Optionally, an automated perfusion system may facilitate delivery of drug compounds to the cellular network in a smooth and controlled manner.

In order to understand more fully the invention, a number of embodiments thereof shall now be described by way of example and with reference to the accompanying drawings,  
30 in which:

Figure 1, is a diagrammatic view of an analysis system in accordance with one aspect of the present invention;

Figure 2, is plan view of a micro-electrode array (MEA) forming part of the system of  
35 Figure 1.

Figure 3, is a flow chart indicative of a method of analysis in accordance with a further aspect of the present invention;

Figure 4 is a diagrammatic view of a sensor in accordance with a further aspect of the invention; and

5 Figure 5 is a graph illustrative of a response to a compound output of the sensor of figure 4.

Figure 6 is a table illustrating a feature set selected for use with the invention;

Figure 7 is a matrix indicative of a set of responses across a feature set for example compounds in accordance with the invention;

10 Figure 8 is a vector diagram based on the responses of Figure 7.

With reference to the Figures and Figure 1, in particular, there is shown an analysis system 1 in which an MEA 3 (see Figure 2) provides a site onto which a cellular network is deposited. The nature of the cellular network is described in respect of a particular  
15 example below, namely a network made up of a plurality of cardio-myocytes. However, it will be appreciated from the outset by those skilled in the art that this particular example of a cellular network is purely exemplary and any reference to cellular network should not be taken to mean purely the example 5 cited herein.

20 The MEA 3 comprises a bio-compatible substrate 5 on which are surface mounted a plurality of electrodes 7 each of which is connected by traces 9 to an edge connector 11. The MEA 3 is insertable within a receptacle 13 formed at the bottom of a well 15. The receptacle 13 is provided with contacts for the edge connector 11. The well 15 is hermetically sealed and forms the environment or chamber 17 within which a compound  
25 is analysed. The chamber is itself connected via a ribbon connector 19 to the input of an amplifier unit 21. The combination of the edge connector 11 and receptacle 13 provides for easy insertion and removal of MEAs 3 for analysis of different compounds. Although Figure 2 illustrates a square array made up of equidistant electrodes, alternative array layouts are contemplated having, for example, non-uniform electrode pitch and layout.  
30 The adoption of a particular layout is predicated on the requirement that an electrode 7 should be capable of detecting electrical activity from a single cell in a network when arranged on the MEA 3. Clearly, the packing or form factor of the MEA 3 must be such that it can be correctly and easily inserted and removed from the receptacle 13 in the well 15.

Output from each electrode 7 passes via the aforementioned cable interconnect 19 to the amplifier unit 21 where the output is amplified. The amplifier unit 21 is a multichannel device capable of providing a gain of around 1000 to each channel. Typically, sufficient channels are available to allow each electrode 7 of the MEA 3 to be mapped to a channel. Depending on the configuration of the MEA 3 there may be need for more or less channels for satisfactory data collection. The amplifier unit 21 itself is interfaced to a PC -based data acquisition system 23. The PC system 23 incorporates an Analogue to Digital conversion card 25 coupled via a PCI bus to a central processor unit. The card 25 provides the external connections necessary to interface the analogue output of the amplifier 21 to the acquisition system 23. The card 25 is capable of sampling the amplified channel data from the amplifier unit 21 at up to 50 kHz/channel. The central processor unit carries out instructions provided by software and/or firmware necessary to analyse the digitised data. The data may be analysed in real time as events occur on the MEA 3 or retrieved later from a storage device 27 such as a hard disk. In the former case, the storage device 27 may still be utilised to archive the data for later repeated or further analysis. The ability to proceed with real time as opposed to or off-line analysis will in some part depend on the rate at which data is generated and the storage capacity of the system 23. The nature of the cellular network placed on the MEA 3 determines the sampling rate. Thus the software and/or firmware is provided with logic to enable the system 23 to operate at an optimum sampling rate for the particular cellular network taking into account any limitations of processor speed and storage capacity. Thus, in the case of a cellular network made up of cardio-myocytes, activity may occur relatively slowly over a 100mS window whereas in the case of neurones, activity may be present in much shorter windows of around 15 to 25mS. In the former case, a relatively lower sampling rate may be adopted by the system 23, with the relevant control signals being provided to the Analogue to Digital Converter 25, than is needed for equivalent resolution of detail in the data derived from a neurone network. The PC system 23 is provided with a VDU 29 and printer for the presentation of results.

In use, and referring especially to Figure 3 again figure of the order of 1000 is applied to each analogue channel which, in the example whose preparation is described below, namely a cardio-myocyte cellular network, has a pre-amplification value of around 100microvolts to 2 millivolts. At this stage, the output from the electrodes 7 on the MEA 3 is an analogue signal. Clearly, before digital signal processing can be applied there is a need to digitise the signal. The rate at which the analogue signal is sampled must be selected to be high enough to ensure that all the features of interest in the electrical

activity of a cellular network deposited on the MEA 3, are available to the digital signal processing unit. As a first step, the signals are amplified 31 and, the analogue signal from each electrode 7 is conditioned 33, in this case by low pass-filtering to remove unwanted high-frequency components. The filtered analogue signal on each channel is  
5 sampled 35 at a rate which may be as much as 50kHz the actual rate 37 depending 39 on the nature of the cellular network placed on the MEA 3. In the case of a cardio-myocyte network, an effective sampling rate has been found to be 10kHz.

By selecting a sampling rate of 10kHz in the example of a cardiac myocyte cellular  
10 network, sufficient resolution is achieved without excessive data collection. For long-term recordings greater than 1 minute in duration, data may be stored as a series of 'cut-outs' of the electrical activity from the cardiac myocyte cells. Each cellular event stored as a 'cut-out' is determined by setting 41 a threshold level (usually a positive value) of at least 2 root-mean-square values of the noise above the baseline. For each event, a time  
15 stamp is recorded at the point at which the threshold level is crossed 43. In addition, electrode raw data 15msec before and 85msec after the threshold level has been crossed is stored. Data is saved to a file on the hard disk in a .mcd file format (approximately 10Mb per electrode for a 60 second recording).

20 The data stored on the hard disk is representative of the changes in electrical activity that occur within a cellular network and are typically in the form of action potentials or spikes. As will be described further below changes occur to the shape of the spikes and their temporal and spatial pattern when a compound is introduced to a cellular network. The electrical activity data from the cellular network is analysed in software by imposing  
25 temporal and spatial information onto a model of the electrode array. In this way, both local and global properties of the electrical activity across the tissue sample can be identified and quantified. Examples of local properties are the peak height, amplitude or depolarisation time of an action potential detected at an individual electrode. Examples of global properties are the beat frequency and propagation speed of action potentials  
30 across the culture. These various properties are referred to as features. The process of feature extraction 47 may then be performed on this digitised data.

In a variant of the present embodiment, before data is stored in a datafile, the features themselves may be used to reduce the requirements for data storage. Thus the spikes  
35 can first be identified using a threshold detector, catalogued and stored, the rest of the



data being ignored. Since the temporal length of a spike is typically much less than the time separation between spikes this procedure requires less storage capacity.

A non-exhaustive set of features identifiable by the analysis system 23 is listed below.

5

#### Examples of Features

**Mean Spike Rate** – the number of spikes observed in a channel divided by the record length. The mean spike rate feature may be updated every minute or few minutes rather than over the whole course of an experiment.

10

**Spike Rate Variability** - calculated from the time differences between consecutive spikes, averaged over all channels.

**Spike Speed** – the propagation speed of the spike pulse across the MEA plate, calculated for each pulse from the spike time arrivals at each selected channel using a least mean squares fit to the data on the assumption of a single plane wave pulse propagating with constant speed.

15

**Arrival Angle** – the direction of propagation of the spike pulse.

20

**Increase in Peak Level** – the relative increase between control and test data in the maximum level of the spike profile averaged over all spikes and all selected channels.

**Increase in Trough Level** – the relative increase between control and test data in the minimum level of the spike profile averaged over all spikes and all selected channels.

25

**Increase in Peak-to-Trough Level** – the relative increase in the range of spike profile averaged over all spikes and all selected channels.

**Increase in Absolute Peak Level** – the relative increase in the maximum absolute level of the spike profile averaged over all spikes and all selected channels.

30

**Increase in Rise Time from 10%** - the relative increase in the time for a spike to achieve its maximum level starting from a level of 10% of the maximum, averaged over all spikes and all selected channels.

5 **Increase in Rise Time from 20%** - the relative increase in the time for a spike to achieve its maximum level starting from a level of 20% of the maximum, averaged over all spikes and all selected channels

**Increase in Recovery Time to 10%** - the relative increase in the time for a spike to recover to 10% of its minimum value starting from the minimum value, averaged over all spikes and all selected channels

10 **Increase in Recovery Time to 20%** - the relative increase in the time for a spike to recover to 20% of its minimum value starting from the minimum value, averaged over all spikes and all selected channels

**Increase in Peak-to-Trough Time** - the relative increase in the time between the maximum level and the minimum level in a spike profile, averaged over all spikes and all  
15 selected channels.

**Increase in Absolute Profile Area** - the relative increase in the area under the modulus profile, averaged over all spikes and all selected channels.

20 **Increase in Profile Rise Area** - the relative increase in the area under the profile between the start and the maximum value, averaged over all spikes and all selected channels.

**Increase in Profile Recovery Area** - the relative increase in the area under the profile  
25 between the minimum value and the end, averaged over all spikes and all selected channels.

**Increase in Absolute Profile Recovery Area** - the relative increase in the area under the modulus profile between the minimum value and the end, averaged over all spikes  
30 and all selected channels.

**Increase in Profile Correlation Coefficient** - the normalised cross-correlation between the control and test spike profiles, averaged over all spikes and all selected channels.

35 **Increase in Profile Variance** - the relative increase in the variance of the spike profile, averaged over all spikes and all selected channels.

**Increase in Profile Skewness** - the relative increase in the skewness of the spike profile, averaged over all spikes and all selected channels.

- 5    **Increase in Profile Kurtosis** - the relative increase in the kurtosis of the spike profile, averaged over all spikes and all selected channels.

**Increase in maximum of wavelet transform** - the relative increase in the maximum value over scale and time delay of the wavelet transform of the spike profile, using for  
10    example a Daubechies wavelet of order 2 here and below, averaged over all spikes and all selected channels.

**Increase in variance of wavelet transform** - the relative increase in the variance of the wavelet transform values of the spike profile, summed over scale and time delay,  
15    averaged over all spikes and all selected channels.

**Wavelet cross-correlation coefficient** - the normalised cross-correlation in scale and time delay between the wavelet transforms of the control and test spike profiles, averaged over all spikes and all selected channels.

20

**Increase in wavelet transform transfer coefficient** – similar to the wavelet cross-correlation coefficient, except that it is normalised by the autocorrelation of the wavelet transform of the control data, instead of by the square root of the product the autocorrelation of the wavelet transform of the control data and the autocorrelation of the  
25    wavelet transform of the test data.

**Increase in profile entropy** - the relative increase in the entropy of the spike profile as determined from its histogram, averaged over all spikes and all selected channels.

- 30    Another feature set which is believed to particularly effective in forming the basis for analysis is set out below and repeated in tabular form as Figure 6 of the drawings. This feature set provides a numerical description of the various changes in the heart beat profile when a drug is applied

35

**Instantaneous Spike Rate** – the relative increase between control and test data in the instantaneous spike rate averaged over all selected channels.

5 **Instantaneous Spike Rate Variability** - the relative increase between control and test data in the temporal variability of the instantaneous spike rate.

**Spike Speed** – the relative increase between control and test data in propagation speed of the spike pulse across the MEA plate, calculated for each pulse from the spike time arrivals recorded at each selected each channel.

10 **Spike Speed Variability** - the relative increase between control and test data in the temporal variability of the spike speed.

15 **Peak Level** – the relative increase between control and test data in the maximum value in the averaged spike profile obtained by averaging the profiles of all the spikes in each selected channel.

20 **Trough Level** – the relative increase between control and test data in the minimum value in the averaged spike profile obtained by averaging the profiles of all the spikes in each selected channel.

**Peak-to-Trough Level** – the relative increase between control and test data in difference between the maximum and minimum values in the averaged spike profile obtained by averaging the profiles of all the spikes in each selected channel.

25 **Absolute Peak Level** – the relative increase between control and test data in the maximum value in the absolute averaged spike profile obtained by averaging the profiles of all the spikes in each selected channel.

30 **Rise Time to 10%** - the increase between control and test data in the time for an averaged spike to achieve its maximum level starting from a level of 10% of the maximum.

35 **Rise Time to 20%** - the increase between control and test data in the time for an averaged spike to achieve its maximum level starting from a level of 20% of the maximum.

**Recovery Time to 10%** - the increase between control and test data in the time for an averaged spike to recover to 10% of its minimum level.

- 5    **Recovery Time to 20%** - the increase between control and test data in the time for an averaged spike to recover to 20% of its minimum level.

**Peak-to-Trough Time** - the increase between control and test data in the time between the maximum level and the minimum level in the averaged spike profile.

10

**QT Time** - the increase between control and test data in the time between the 3% and 97% cumulative points of the absolute area under the averaged spike profile.

- 15    **Profile decay rate** - the increase between control and test data in the decay rate of the tail of the averaged profile.

**Absolute Profile Area** - the relative increase between control and test data in the area under the modulus of the averaged profile.

- 20    **Profile Rise Area** - the relative increase between control and test data in the area under the averaged profile between the start and the maximum value.

- 25    **Absolute Profile Recovery Area** - the relative increase between control and test data in the area under the modulus of the averaged profile between the minimum value and the end.

**Profile turning moment** - the relative increase between control and test data of the temporal turning moment of the averaged profile.

- 30    **Absolute profile centre of gravity** - the relative increase between control and test data of the centre of gravity of the absolute averaged profile.

- 35    **Absolute profile radius of gyration** - the relative increase between control and test data of the radius of gyration of the absolute averaged profile measured about its centre of gravity.

**Amplitude Variance** - the relative increase between control and test data in the variance of the averaged spike profile.

5 **Maximum spectral value** - the relative increase between control and test data in the maximum value of the power spectrum of the averaged spike profile.

**Frequency of maximum spectral value** - the relative increase between control and test data in the frequency of the maximum value of the power spectrum of the averaged spike profile.

10

**Amplitude Variance in Frequency Band 1** - the relative increase between control and test data in the variance of the averaged spike profile, normalised by the total variance, in the frequency band 0-250 Hz.

15 **Amplitude Variance in Frequency Band 2** - the relative increase between control and test data in the variance of the averaged spike profile, normalised by the total variance, in the frequency band 250-500 Hz.

20 **Amplitude Variance in Frequency Band 3** - the relative increase between control and test data in the variance of the averaged spike profile, normalised by the total variance, in the frequency band 500-750 Hz.

25 **Amplitude Variance in Frequency Band 4** - the relative increase between control and test data in the variance of the averaged spike profile, normalised by the total variance, in the frequency band 750-1000 Hz.

**Amplitude Variance in Band 2 / Band 1** - the relative increase between control and test data in the ratio of the variances in bands 2 and 1 in the spectrum of the averaged spike profile.

30

**Amplitude Variance in Band 3 / Band 2** - the relative increase between control and test data in the ratio of the variances in bands 3 and 2 in the spectrum of the averaged spike profile.

35 **Amplitude Correlation Coefficient** - the normalised cross-correlation between the averaged control and averaged test spike profiles.

**Amplitude Skewness (normalised)** - the relative increase between control and test data in the skewness, normalised with respect to the total variance, of the averaged spike profile.

5

**Amplitude Kurtosis (normalised)** - the relative increase between control and test data in the kurtosis, normalised with respect to the total variance, of the averaged spike profile.

10 **Entropy of profile** - the relative increase between control and test data in the entropy of the averaged spike profile as determined from its histogram.

**Entropy of absolute profile** - the relative increase between control and test data in the entropy of the absolute averaged spike profile as determined from its histogram.

15

**Maximum wavelet transform coefficient** - the relative increase between control and test data in the maximum value over scale and time delay of the wavelet transform of the averaged spike profile, using a Daubechies wavelet of order 2 here and below.

20 **Scale change of wavelet transform coefficient** - the relative increase between control and test data in the scale of maximum value over scale and time delay of the wavelet transform of the averaged spike profile, using a Daubechies wavelet of order 2 here and below.

25 **Variance of wavelet transform** - the relative increase between control and test data in the variance of the wavelet transform values of the averaged spike profile, summed over scale and time delay.

**Wavelet transform transfer coefficient** - wavelet cross-correlation coefficient  
30 normalised by the autocorrelation of the wavelet transform of the control data.

**Variance of wavelet transform ridge values** - the relative increase between control and test data in the variance of the vector of wavelet transform values of the averaged spike profile obtained by taking the maximum value at each scale.

35

**Wavelet transform transfer coefficient ridge values** – wavelet cross-correlation coefficient of the maximum vector as defined above, normalised by the autocorrelation of the corresponding vector in the control data.

5

It should be noted that not all the above features are amplitude dependant. Thus, features which depend on the recovery rate of the cellular network may be used to assist in detection and classification. Furthermore, although the above features may or may not be present to a greater or lesser extent in the electrical activity of the cellular network it is considered that similar features may be identified from the chemical behaviour of the network in respect of its fluorescent and/or luminescent activity.

10

Figure 7, exemplifies the results of a feature set analysis in a matrix format for a set of different compounds aetyl choline I, cafeine II, pinacidil III, salubutamol IV and drug C V, such as might result from carrying out activities set out in the examples below. The numerals 100 beneath the columns are indicative of respective features in a feature set and the level of shading of the boxes indicates the nature of the response. Results from a control VI are shown separately.

15

Figure 8 shows the outcome of plotting the results of a feature set as a vector quantity in three-dimensional space for the compounds of Figure 7 where clustering of the results is evident for each of the compounds identified by their respective reference numerals.

20

The sequence of activities necessary to analyse a compound is set out below. These activities are a combination of physical processes taken in relation to the deposition of a cellular network on the MEA and the compound to be tested together with signal processing activity which is carried out utilising the PC-based system 23. Firstly a cellular network is cultured and deposited on the MEA 3. An example of the steps required in this regard using cardio-myocyte cells is as follows:

25

30

### **Example**

(a) heart tissue is isolated from rat embryos (E15-E18) or neonates.

(b) The heart tissue is minced using a scalpel and placed into cold (4°C)  $\text{Ca}^{2+}/\text{Mg}^{2+}$  - free Hanks balanced salt solution (HBSS)

35



- (c) The tissue is washed (3 times) in fresh HBSS and replaced with 0.05% trypsin in HBSS
- (d) Incubate 10 mins at 37°C and discard supernatant
- (e) Fresh DNase type II solution added (10,000 Units/ml) for 2 mins
- 5 (f) Fresh trypsin added and incubated at 37°C for 8 mins
- (g) Supernatant removed and placed into HAMS F10 containing 36% FCS, 0.5% insulin/transferrin/selenite, 6mM L-glutamine and 2% penicillin/streptomycin (200mM stock)
- (h) Cells are collected from suspension (1500rpm, 5mins) and resuspended in HAMS
- 10 F10 containing 10% FCS, 0.5% insulin/transferrin/selenite, 6mM L-glutamine and 2% penicillin/streptomycin (200mM stock)
- (i) Steps (e) – (h) repeated 5-8 times
- (j) Differential adhesion is performed by incubating pooled cell suspensions in a treated tissue culture flask for 2 hrs at 37°C
- 15 (k) The final cell suspension is counted and seeded onto fibronectin treated MEA plates at 50K per plate in a 100µl volume.

The cell suspension is deposited on the MEA 3 such that each electrode 7 is in contact with a respective cell.

20

Once the cellular network is in place on the MEA 3, the MEA 3 is inserted within the receptacle 13 in the well 15 such that a set of baseline measurements may be recorded. Accordingly, the analogue output from the electrodes 7 is amplified 31, filtered 33 and stored 45 utilising the equipment and methodology set out above. Throughout this

25 baseline assessment stage and subsequently during the analysis of a compound, it is desirable to maintain the cell culture conditions substantially constant. Thus, the environment is monitored utilising sensors deployed in the cell culture chamber 17 encompassing the MEA. The sensor output is monitored by a software module 18 running on the PC-based data acquisition system 23 or it may be monitored

30 independently. In either case, the signals received from the sensor are employed to adjust environmental controls. The parameters which may be monitored could include temperature, pH and dissolved oxygen concentration of the culture medium.

The recording process for the baseline measurements is performed for each data

35 channel corresponding to an electrode 7. For a given record (typically 100 seconds long)

a set of so-called healthy channels is identified as follows by identifying the set of channels with the most frequently occurring non-zero number of spikes.

5 A compound to be analysed is then introduced to the cellular network covering the MEA 3. This may be applied directly to the network on the MEA 3 used for baseline assessment or a further equivalent sample of cultured network may be prepared on a further MEA 3 and inserted in the receptacle 13 in the well 15 in its place. Again, for a given record (typically 100 seconds in duration although a longer or shorter period may be selected) a set of so-called healthy channels is selected by identifying the set of  
10 channels with the most frequently occurring non-zero number of spikes. The healthy channels in each set i.e. the baseline measurements and the subsequent measurements in the presence of the compound are then compared further to identify the channels which are common to both sets. These common channels are then subjected to feature extraction to form a feature set.

15 To be useful in the subsequent detection and classification stages, a feature must be readily extractable from the data and numerically quantifiable. Various processing algorithms are used to extract features meeting this requirement. As many such features as possible are included in the set to encompass as much of the information content of the data as possible. The presence of redundant features in the set is tolerated. Furthermore, by averaging as late as possible in the process, the sensitivity of the features is found to be enhanced. The significance of a measured feature can also be estimated by calculating the standard deviation of that measure over all the selected  
20 channels.

25 The feature set can then be viewed as a vector quantity, with dimensions equal to the number of features; each component representing the numerical change to the feature in question equal to the difference between the baseline and subsequent measurements. Detection and classification reduces to performing manipulations on the response vectors. The detection process has been described above. The classification process is achieved by the use of standard cluster analysis by which is to be understood those mathematical clustering and partitioning techniques which can be used to group cases on the basis of their similarity over a range of variables e.g. component. Many clustering algorithms are available; they differ with respect to the method used to measure  
30 similarities (or dissimilarities) and the points between which distances are measured. Thus, although clustering algorithms are objective, there is scope for subjectivity in the

selection of an algorithm. The most common clustering algorithms are polythetic agglomerative, i.e. a series of increasingly larger clusters are formed by the fusion of smaller clusters based on more than one variable. This hierarchical approach is particularly suited to computer based analysis in view of the large data sets which are to be analysed. However, a less computationally intensive and therefore more rapid approach is the non-hierarchical k means, or iterative relocation algorithm. Each case is initially placed in one of  $k$  clusters, cases are then moved between clusters if it minimises the differences between cases within a cluster. Depending on the computational capability of the PC-acquisition system and subject to any requirement for real-time analysis one of the aforementioned processes may be adopted.

In addition to the embodiment set out above, further variants having different MEA 3 arrangements are contemplated, some of which are set out below:

- (1) SINGLE WELL, MULTIPLE CHAMBER ENVIRONMENT MEASURES – has sensors incorporated into the chamber apparatus to include simultaneous measurement and control of the culture environment (controlled perfusion, temperature, pH sensors).
- (2) SINGLE WELL, MULTIPLE CHAMBER ENVIRONMENT MEASURES, MULTIPLE CELL PHYSIOLOGY MEASURES – as in (1) with integrated sensors to enable measurement of other cell functions such as intracellular calcium levels, lactate production etc.
- (3) MULTI-WELL, MULTI-CHANNEL SYSTEM – instead of a single well format, 96wells are formed and data retrieved from multiple microelectrodes in each well.
- (4) MULTI-WELL, MULTI-PARAMETER SYSTEM – combination of (1) and (3) producing a drug screening device. A completely controlled multiwell assay system capable of controlled delivery of drug and fully automated data capture and analysis
- (5) MULTI-WELL, MULTI-PARAMETER SYSTEM, MULTIPLE CELL PHYSIOLOGY MEASURES. Combination of (2) and (3) to allow integrated analysis of many cell functions in many different assays.

In accordance with a further embodiment of the invention, there is provided a sensor for the detection of compounds such as environmental toxins. Figure 4 shows a schematic diagram of the sensor in which the cellular network is provided by a culture of cells derived from a species that is subject to a threat or a species whose response to a toxin, for example, may be extrapolated to another species such as humans. For example, the cellular network may comprise scallop heart cells.

The sensor 101 includes a plurality of chambers 117 each capable of housing an MEA 103. A perfusion system 118 permits the delivery of samples to be tested, for example a test sample derived from river water, to each MEA 103. The MEA 103 is as previously  
5 described in relation to the first embodiment in that it has a plurality of electrodes 107 arranged to contact the cellular network, in use. Electrical signals output from the electrodes 107 pass via an interconnect 119 to the amplifier unit 121 where the output is amplified. The amplifier unit 121 is a multichannel device capable of providing a gain of around 1000 to each channel. Typically, sufficient channels are available to allow each  
10 electrode 107 of the MEA 103 to be mapped to a channel. Depending on the configuration of the MEA 103 there may be need for more or less channels for satisfactory data collection. The amplifier unit itself is interfaced to a PC -based data acquisition system 123. The PC system 123 incorporates an Analogue to Digital conversion card 125 coupled via PCI bus to a central processor unit 126. The card 125  
15 provides the external connections necessary to interface the analogue output of the amplifier 121 to the acquisition system 123. The card 125 is capable of sampling the amplified channel data from the amplifier unit 121 at up to 50 kHz/channel. The actual rate is determined by reference to the nature of cellular network and the resolution required to identify features of interest. The central processor unit 126 carries out  
20 instructions provided by software and/or firmware necessary to analyse the digital data. The data may be analysed in real time as events occur on the MEA 103 or retrieved from a storage device 127 such as a hard disk. In the former case, the storage device 127 may still be utilised to archive the data for later repeated or further analysis. The ability to proceed with real time as opposed to or off-line analysis will in some part depend on  
25 the rate at which data is generated and the storage capacity of the system 123.

The software and/or firmware includes a library 128 of feature sets corresponding to experimental or otherwise derived responses to a particular toxin or toxins for a particular cellular network under threat. As such, the sensor 101 typically is required to perform an  
30 identification and preferably classification function in relation to a pre-defined set of feature sets, i.e. there is no requirement for the sensor 101 to identify every compound that is present in the sample only those whose effect might be toxic to the species under threat. Depending on the particular species under investigation, different libraries of feature sets may be loaded into the sensor 101. Conveniently, the sensor 101 is  
35 provided with software specific to the species with which a user is intending to work. Such software will contain the libraries as integrated portions of the software or as user

loadable software modules. For example, accumulation of toxins in shellfish represents a growing problem. Two toxins are of particular importance because of their profound effects on the human nervous system. These are amnesiac shellfish poisoning (ASP) and paralytic shellfish poisoning (PSP) toxins. Ingestion of large doses of these toxins  
5 can result in death.

Taking ASP poisoning as an example of the use of the sensor 101, it is known that the toxin primarily involved in ASP is domoic acid. Accordingly, before the sensor 101 may be utilised to detect the presence or otherwise of a particular toxin, it is necessary to  
10 provide a feature set for inclusion in the library 128 for subsequent use in the sensor. The creation of such a feature set or vector can be carried out using the system – described in the first embodiment. In the case of ASP, as has been indicated, the primary toxin is domoic acid. Thus, a network of cortical neurone cells from a mouse is applied to an MEA. The MEA is then placed within the well 115 and as has been  
15 described previously, the electrical response of the network to the addition of the compound, in this case domoic acid is extracted by the electrodes 107, amplified and the active channels identified and the relevant data captured. This data is then analysed, again as has been described above and in the case of domoic acid it is found that one particularly useful feature indicative of domoic acid is Mean Spike Rate (MSR). Figure 5  
20 is a graph illustrating the mean spike rate from rat cortical neurone culture when exposed to domoic acid of 100nM concentration at 10 minutes from the start of the data as indicated by the heavy line T. This response is then parameterised and stored in a file for inclusion in a software library of features.

Subsequently, a sensor 101 is used to determine the presence or otherwise of ASP in a  
25 sample of shellfish. The sensor 101 incorporates storage for a library of feature sets or vectors indicative of the responses to particular toxins realised by particular species. The library may be downloaded to the sensor such that it is held on a hard disk or other storage media 127 or it may be accessed via a network connection to a database. In one particular variant, the library is stored in an integrated circuit formed on the MEA 103  
30 itself. Thus, by means of colour coding or another indicia, an appropriate MEA 103 is selected having the requisite prestored library of feature sets or vectors applicable to a particular species under study. The integrated circuit is provided with the appropriate connections to the card edge connector to allow the library to be accessed by the sensor when installed in the receptacle in the well 115.

Either before or after installation in the receptacle 113, a cellular network of rat derived cortical neurone cellular material is deposited on an MEA 103. The MEA 103 is placed within the receptacle 113 in the well 115 of the sensor 101 and a shellfish sample is placed upon the MEA 103. The electrical response of the network to the addition of the compound (in this case the shellfish sample) is extracted by the electrodes 107, amplified and the active channels identified and the relevant data captured. This data is then analysed against a library of feature sets including that obtained for ASP in relation to the MSR feature. If a match with a library response is found, within a predetermined limit of statistical confidence, then a positive result is flagged and an appropriate warning indication is given by the sensor 101 which may be visual, audio or a combination thereof. Where no such response is identified, again with a particular statistical confidence, then no such warning is given merely an indication with a level of confidence that the sample seems to be toxin free. Clearly, further features may exist which are found to correlate strongly with the presence or otherwise of a particular toxin. Such a feature set may be utilised in the analysis of unknown compounds and a match with one or more may be sufficient to cause the sensor to generate a warning indication.

In a further embodiment of the invention, the PC system 23 may be utilised as a measuring tool. Thus, the system 23 may be utilised to assess the physical and/or chemical characteristics of a known compound. For example, it has been recognised that the response of the system 23 to a particular compound may depend on the concentration of that compound. By establishing a library of features or feature set indicative of particular concentration levels at predetermined levels of statistical confidence, the concentration of a particular known compound may be identified. It will be further appreciated that the measurement tool may be integrated with the identification and classification function such that an initially unknown compound may be both identified and particularised in terms of its physical/chemical properties.

It will be recognised by those skilled in the art that a number of factors will impinge on the statistical confidence a user may have in the results of analysis using the method and apparatus of the invention. A quantification of the impact such factors might have may be incorporated in the statistical level of confidence applied to the results of a particular analysis. In order to minimise such effects, steps may be taken to utilise features that are normalised relative to a control. Alternatively, the minimum concentration to produce a significant response could be employed. Additionally, the response of the system may

depend on the proximity of each cell to its nearest electrode. This may require the use of features that are independent of absolute amplitude, such as beat rate.

## Claims

1. A compound analysis method, the method comprising determining a vector quantity having a number of dimensions equal to a number of features derived from the electrical output of an electrically active cellular network with each component of said vector being representative of a change in said feature resulting from the application of a compound to said electrically active cellular network and classifying said vector in accordance with a predetermined cluster analysis methodology.
2. A method as claimed in Claim 1, including providing a library of features characterising known compounds such that classification of said vector enables identification of the compound in accordance with a predetermined measure of statistical reliability.
3. A compound analysis system, the system comprising a micro-electrode array provided by a bio-compatible substrate having a plurality of electrodes situated thereon, said electrodes having an arrangement on said substrate corresponding substantially to that of an electrically active cellular network disposable in use thereon, a multi-channel amplifier coupled to said electrodes and an analyser operatively connected to said amplifier to determine for each active channel a vector quantity having a number of dimensions equal to a number of features derived from the electrical output of said electrically active cellular network with each component of said vector being representative of a change in said feature.
4. A compound analysis apparatus, the apparatus including a processor and a memory, the processor being operable in response to signals derived from a micro-electrode array connected, in use, thereto, to determine a vector quantity having a number of dimensions equal to a number of features derived from the electrical output of said micro-electrode array with each component of said vector being representative of a change in said feature, wherein said memory contains a library of features characterising known compounds such that classification of said vector enables identification of a compound deposited, in use, on said array, in accordance with a predetermined measure of statistical reliability.



5. An apparatus as claimed in Claim 4, including a storage device, such that signals derived from said array are held by said storage device.
6. A sensor for compound detection, the sensor comprising a receptacle for a micro-electrode array, said receptacle having a connector for receiving electrical signals from said array when received in said receptacle, an amplifier for amplifying said signals and a processor, the processor being operable in response to said signals to determine a vector quantity having a number of dimensions equal to a number of features derived from the electrical output of said micro-electrode array with each component of said vector being representative of a change in said feature.
7. A sensor as claimed in Claim 6, further including a memory, said memory containing a library of features characterising known compounds such that classification of said vector enables identification of the compound deposited, in use, on said array,
8. A sensor as claimed in Claim 6 or Claim 7, wherein the memory is integral therewith.
9. A micro-electrode array for use in compound analysis, the array comprising a bio-compatible substrate having a plurality of electrodes situated thereon, said electrodes being coupled to a connector, said connector providing, in use, electrical connectivity to an analysis apparatus, wherein the array further includes a memory coupled to said connector, said memory being accessible, in use, by said analysis apparatus.
10. A computer program product in a computer-readable medium for use in a compound analysis apparatus, the computer program product comprising determining a vector quantity having a number of dimensions equal to a number of features derived from an electrical output of a micro-electrode array with each component of said vector being representative of a change in said feature, accessing a memory containing a library of features characterising known compounds and classifying said vector in order to identify a compound deposited, in use, on said array.

11. A computer program product as claimed in Claim 10, including determining a level of statistical reliability applicable to said identification of said compound.
- 5 12. A compound analysis system, substantially as hereinbefore described with reference to Figure 1 of the accompanying drawings.
13. A sensor for compound detection, substantially as hereinbefore described with reference to Figure 4 of the accompanying drawings.

Fig. 1

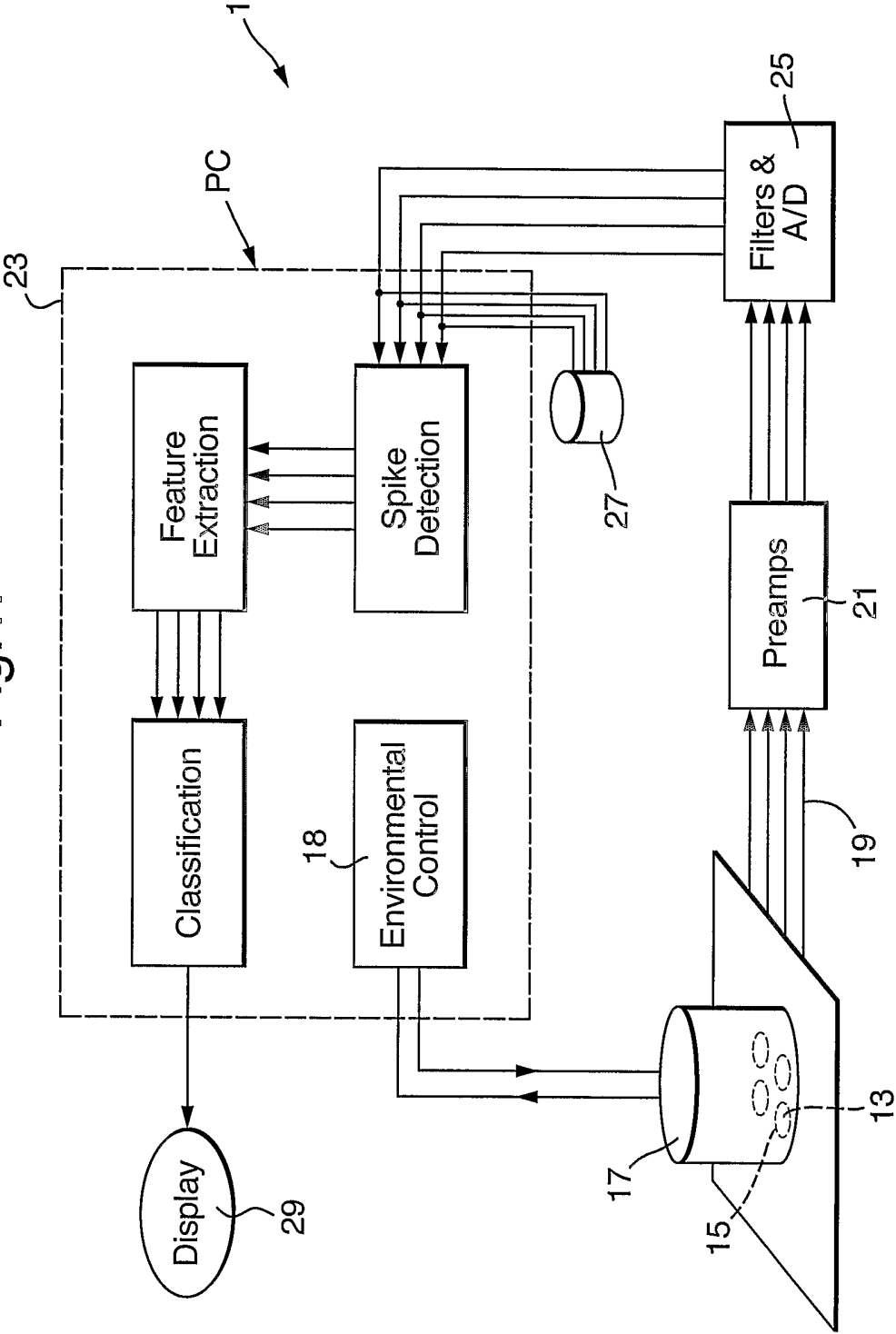
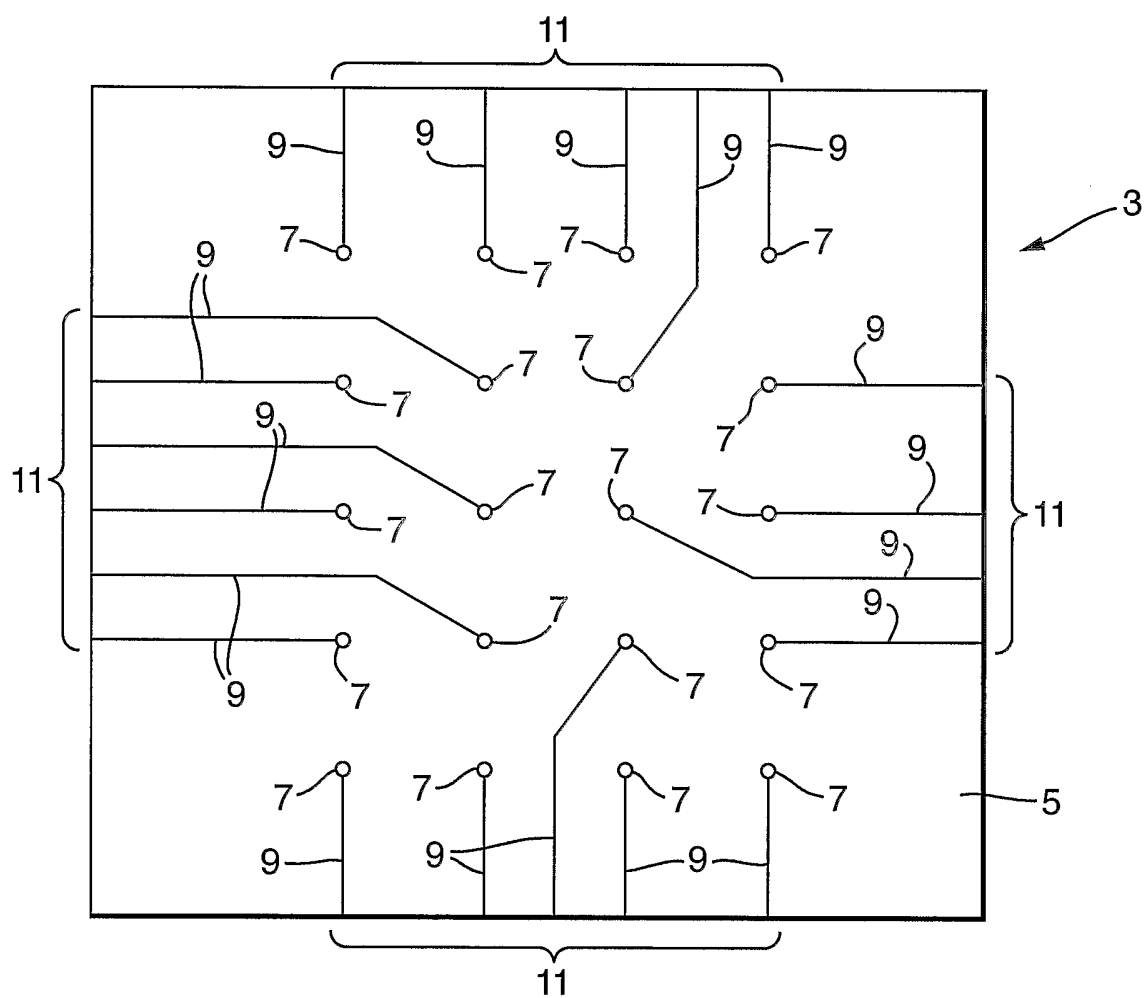


Fig.2.



3/8

Fig.3.

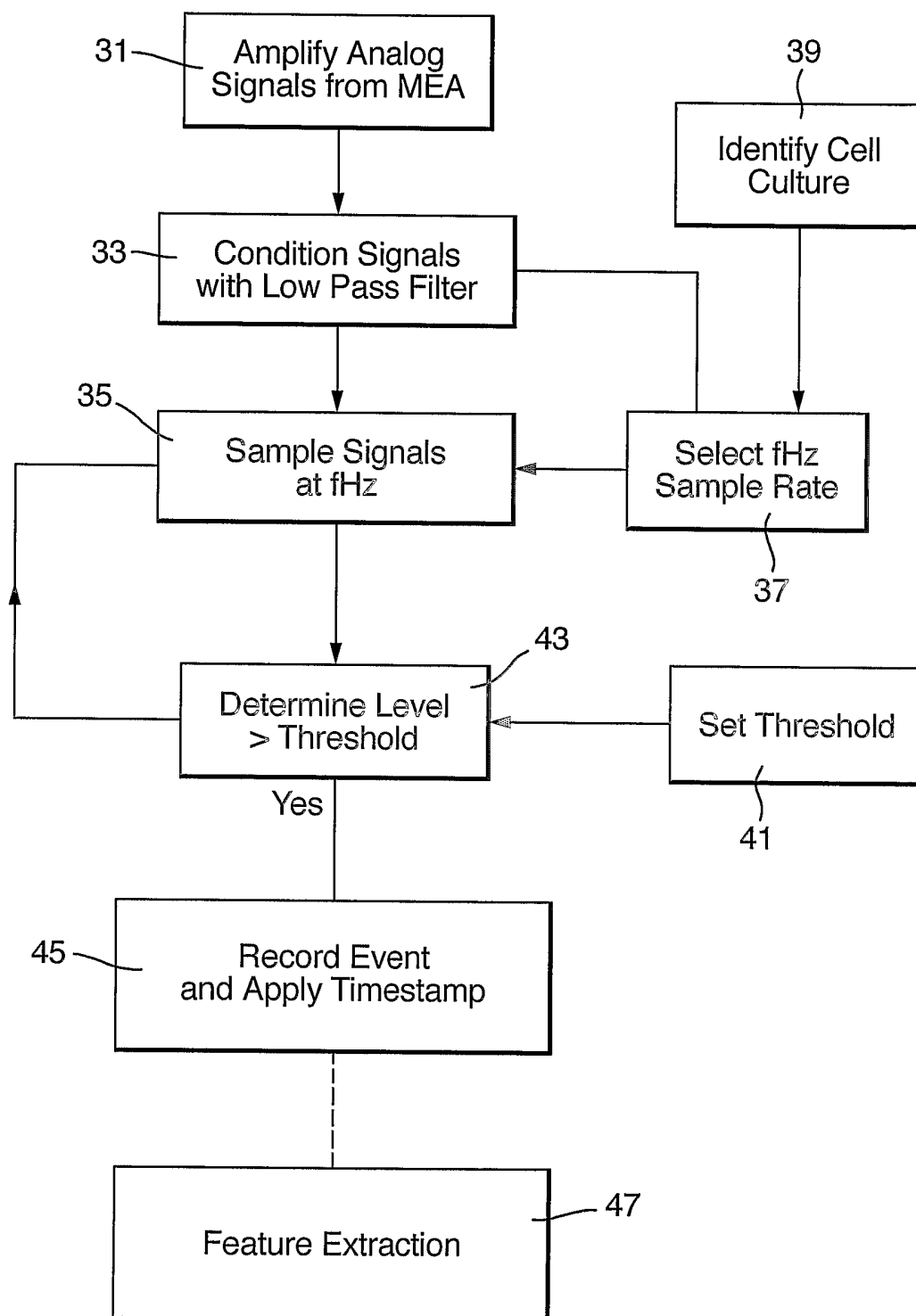


Fig. 4

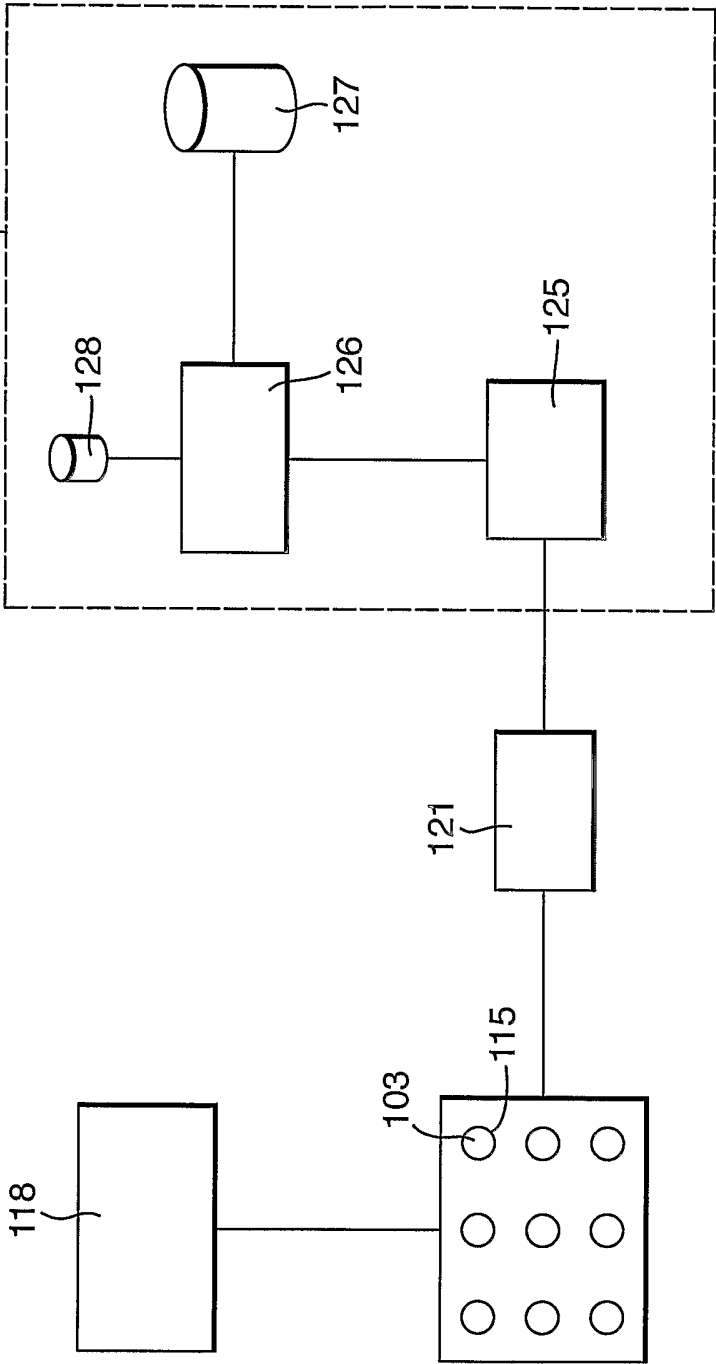


Fig.5<sub>a</sub>

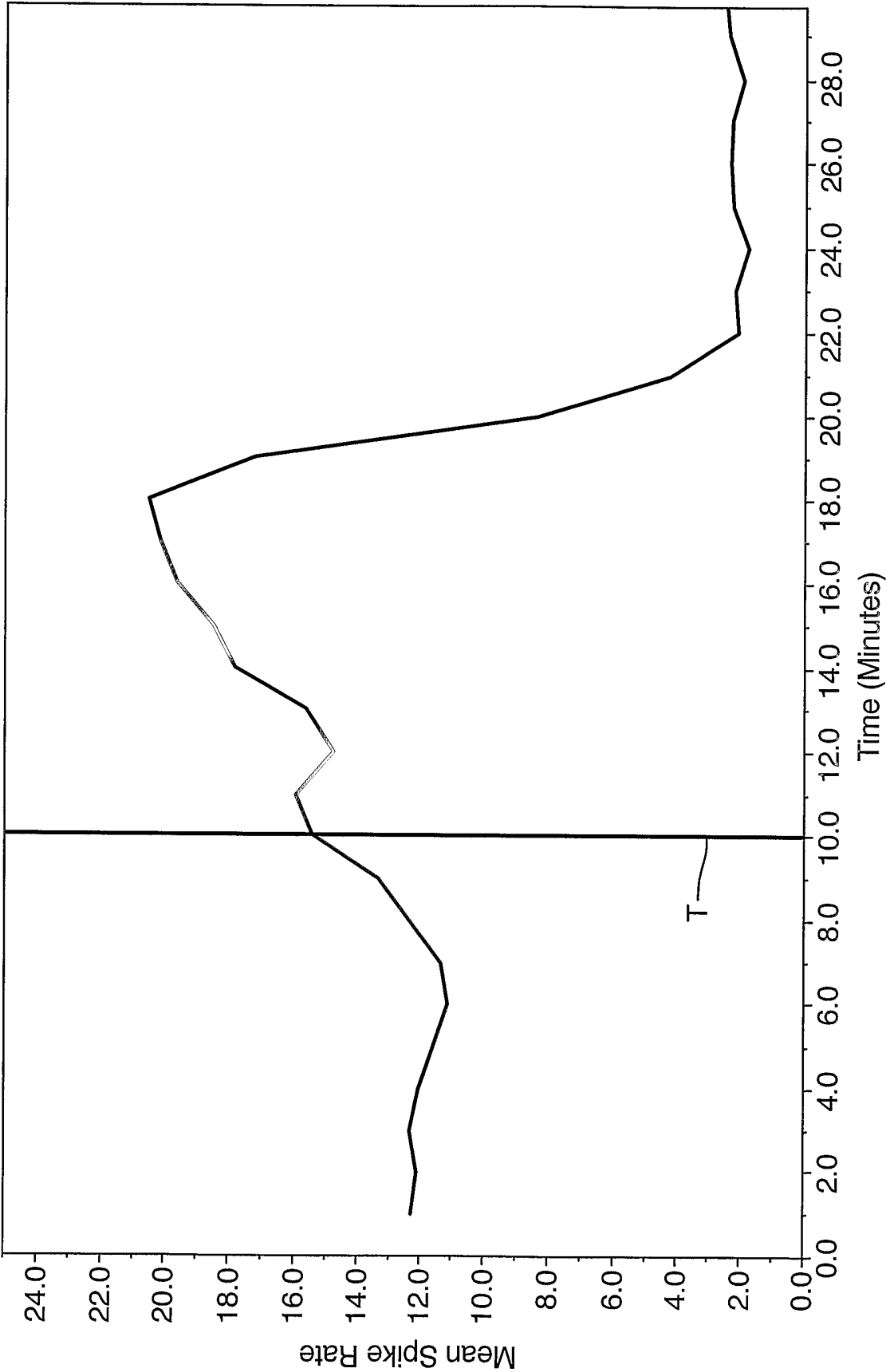
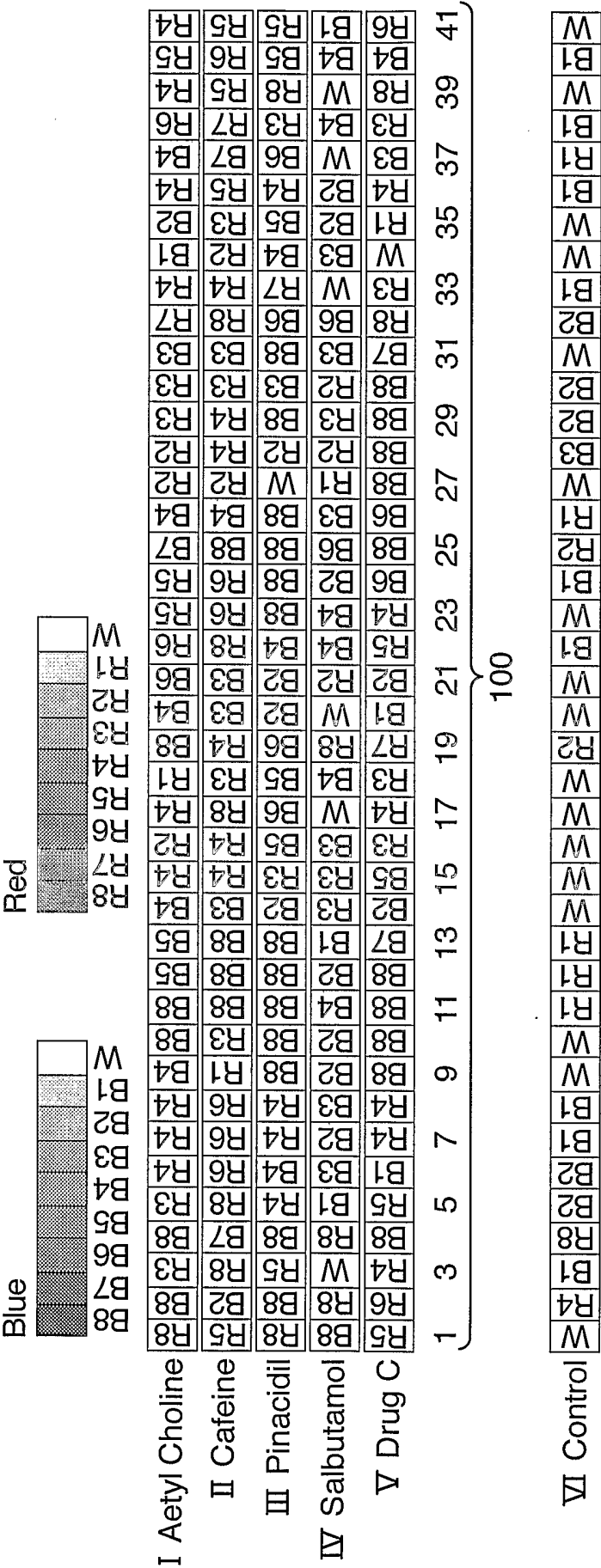


Fig.6.

Program	Feature No.	Name	Description	Type of t-test
spike_rate	1	isr	instantaneous spike rate	unpaired
spike_rate	2	isrv	instantaneous spike rate variability	unpaired
spike_speed	3	ss	spike speed	unpaired
spike_speed	4	ssv	spike speed variability	unpaired
max_min	5	maxv	peak level	paired
max_min	6	minv	trough level	paired
max_min	7	rangev	peak-to-trough level	paired
max_min	8	maxmodv	absolute peak level	paired
rise_times	9	rise10_time	rise time from 10% of peak	paired
rise_times	10	rise20_time	rise time from 20% of peak	paired
rise_times	11	recov10_time	recovery time to 10% of peak	paired
rise_times	12	recov20_time	recovery time to 20% of peak	paired
rise_times	13	diffv	peak-to-trough time	paired
rise_times	14	qt	time from 3% and 97% of absolute profile area	paired
rise_times	15	dec_mpt	decay rate of profile tail	paired
spike_area	16	mean_abs	absolute profile area	paired
spike_area	17	mean_rise	profile rise area	paired
spike_area	18	mean_abs_recov	absolute profile recovery area	paired
moments	19	mom	turning moment of profile	paired
moments	20	cog	centre of gravity of absolute profile	paired
moments	21	rog	radius of gyration of absolute profile	paired
spectrum	22	varv	amplitude variance	paired
spectrum	23	maxspec	maximum power spectral value	paired
spectrum	24	locmaxspec	frequency of maximum power spectral value	paired
spectrum	25	spec1	amplitude variance (normalised) in band 1	paired
spectrum	26	spec2	amplitude variance (normalised) in band 2	paired
spectrum	27	spec3	amplitude variance (normalised) in band 3	paired
spectrum	28	spec4	amplitude variance (normalised) in band 4	paired
spectrum	29	spec2re1	amplitude variance band 2 relative to band 1	paired
spectrum	30	swpec3re2	amplitude variance band 3 relative to band 2	paired
spike_hos	31	ccv	amplitude correlation coefficient	1-sided
spike_hos	32	skew	amplitude skewness (normalised)	paired
spike_hos	33	kurt	amplitude kurtosis (normalised)	paired
entropy	34	ent	entropy of profile	paired
entropy	35	enta	entropy of absolute profile	paired
wavelets	36	wmaxv	maximum wavelet transform coefficient	paired
wavelets	37	scale_maxmaxv	scale of wavelet transform coefficient	paired
wavelets	38	wvv	variance of wavelet transform	paired
wavelets	39	wtrfv	wavelet transform transfer coefficient	unpaired
wavelets	40	wvmv	variance of wavelet transform ridge values	paired
wavelets	41	wtrfmv	wavelet transform ridge value transfer coefficient	unpaired



Fig. 7



8/8

Fig.8.

