

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
8 June 2006 (08.06.2006)

PCT

(10) International Publication Number  
**WO 2006/060200 A1**

(51) International Patent Classification:  
**G06F 19/00** (2006.01)

(21) International Application Number:  
PCT/US2005/042003

(22) International Filing Date:  
18 November 2005 (18.11.2005)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
11/000,681 30 November 2004 (30.11.2004) US

(71) Applicant (for all designated States except US): **AGILENT TECHNOLOGIES, INC.** [US/US]; 395 Page Mill Road, Palo Alto, CA 94306 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **WEBB, Peter, G.** [AU/US]; 345 Willow Road, Menlo Park, CA 94025 (US). **NELSON, Charles, F.** [US/US]; 416 Portofino Drive, #202, San Carlos, CA 94070 (US). **SHUKLA, Amitabh** [US/US]; 5447 Blackoak Way, San Jose, CA 95129 (US). **MUNSON, Marilyn** [US/US]; 105 Vendola Drive, San Rafael, CA 94903 (US).

(74) Agent: **REES, Dianne, M.**; Agilent Technologies, INC., Legal Department DL 429, Ip Administration, P.O. Box 7599, Loveland, CO 80537-0599 (US).

(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, KM, KN, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, 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, NA, 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, IS, IT, LT, LU, LV, 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).

**Published:**

- with international search report
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments

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: SYSTEMS AND METHODS FOR PRODUCING CHEMICAL ARRAY LAYOUTS

(57) Abstract: Systems and methods for using the same to obtain a chemical array layout are provided. Also provided are computer program products for executing the subject methods.



WO 2006/060200 A1

## SYSTEMS AND METHODS FOR PRODUCING CHEMICAL ARRAY LAYOUTS

### BACKGROUND OF THE INVENTION

5 Polynucleotide arrays (such as DNA or RNA arrays) are known and are used, for example, as diagnostic or screening tools. Such arrays include regions of usually different sequence polynucleotides arranged in a predetermined configuration on a substrate. These regions (sometimes referenced as “features”) are positioned at respective locations (“addresses”) on the substrate.

10 The arrays, when exposed to a sample, will exhibit an observed binding pattern. This binding pattern can be detected upon interrogating the array. For example all polynucleotide targets (for example, DNA) in the sample can be labeled with a suitable label (such as a fluorescent compound), and the fluorescence pattern on the array accurately observed following exposure to the sample. Assuming that

15 the different sequence polynucleotides were correctly deposited in accordance with the predetermined configuration, then the observed binding pattern will be indicative of the presence and/or concentration of one or more polynucleotide components of the sample.

Biopolymer arrays can be fabricated by depositing previously obtained

20 biopolymers (such as from synthesis or natural sources) onto a substrate, or by in situ synthesis methods. Methods of depositing obtained biopolymers include loading then touching a pin or capillary to a surface, such as described in US 5,807,522 or deposition by firing from a pulse jet such as an inkjet head, such as described in PCT publications WO 95/25116 and WO 98/41531, and elsewhere.

25 Such a deposition method can be regarded as forming each feature by one cycle of attachment (that is, there is only one cycle at each feature during which the previously obtained biopolymer is attached to the substrate). For in situ fabrication methods, multiple different reagent droplets are deposited by pulse jet or other means at a given target location in order to form the final feature (hence

30 a probe of the feature is synthesized on the array substrate). The in situ fabrication methods include those described in US 5,449,754 for synthesizing peptide arrays, and described in WO 98/41531 and the references cited therein

for polynucleotides, and may also use pulse jets for depositing reagents. The in situ method for fabricating a polynucleotide array typically follows, at each of the multiple different addresses at which features are to be formed, the same conventional iterative sequence used in forming polynucleotides from nucleoside reagents on a support by means of known chemistry. This iterative sequence can be considered as multiple ones of the following attachment cycle at each feature to be formed: (a) coupling a selected nucleoside (a monomeric unit) through a phosphite linkage to a functionalized support in the first iteration, or a nucleoside bound to the substrate (i.e. the nucleoside-modified substrate) in subsequent iterations; (b) optionally, but preferably, blocking unreacted hydroxyl groups on the substrate bound nucleoside; (c) oxidizing the phosphite linkage of step (a) to form a phosphate linkage; and (d) removing the protecting group ("deprotection") from the now substrate bound nucleoside coupled in step (a), to generate a reactive site for the next cycle of these steps. The functionalized support (in the first cycle) or deprotected coupled nucleoside (in subsequent cycles) provides a substrate bound moiety with a linking group for forming the phosphite linkage with a next nucleoside to be coupled in step (a). Final deprotection of nucleoside bases can be accomplished using alkaline conditions such as ammonium hydroxide, in a known manner. Conventionally, a single pulse jet or other deposition unit is assigned to deposit a single monomeric unit.

The foregoing chemistry of the synthesis of polynucleotides is described in detail, for example, in Caruthers, *Science* 230: 281-285, 1985; Itakura et al., *Ann. Rev. Biochem.* 53: 323-356; Hunkapillar et al., *Nature* 310: 105-110, 1984; and in "Synthesis of Oligonucleotide Derivatives in Design and Targeted Reaction of Oligonucleotide Derivatives", CRC Press, Boca Raton, Fla., pages 100 et seq., US 4,458,066, US 4,500,707, US 5,153,319, US 5,869,643, EP 0294196, and elsewhere. The phosphoramidite and phosphite triester approaches are most broadly used, but other approaches include the phosphodiester approach, the phosphotriester approach and the H-phosphonate approach. The substrates are typically functionalized to bond to the first deposited monomer. Suitable techniques for functionalizing substrates with such linking moieties are described, for example, in Southern, E. M., Maskos, U. and Elder, J.K., *Genomics*, 13, 1007-1017, 1992.

In the case of array fabrication, different monomers may be deposited at different addresses on the substrate during any one cycle so that the different features of the completed array will have different desired biopolymer sequences. One or more intermediate further steps may be required in each cycle, such as  
5 the conventional oxidation and washing steps in the case of in situ fabrication of polynucleotide arrays.

In array fabrication, the quantities of polynucleotide available are usually very small and expensive. Additionally, sample quantities available for testing are usually also very small and it is therefore desirable to simultaneously test the  
10 same sample against a large number of different probes on an array. These conditions require use of arrays with large numbers of very small, closely spaced features. A typical array may contain thousands of features. It is important in such arrays that features actually be present, that they are put down as accurately as possible in the desired target pattern, are of the correct size, and  
15 that the probe nucleic acid is uniformly coated within the feature. If any of these conditions are not met within a reasonable tolerance, the results obtained from a given array may be unreliable and misleading. This of course can have serious consequences to diagnostic, screening, gene expression analysis or other purposes for which the array is being used.

Thus, fabricating a required number of arrays, particularly with very high number of features, is often not a task an end user wishes to perform herself. As a result, array users or other customers may turn to specialized array fabricators. As the use of specialized array fabricators grows, there is continued interest in the development of improved methods performing one or more aspects of the  
20 interaction between a customer and array fabricator.  
25

### SUMMARY OF THE INVENTION

Systems and methods for using the same to obtain a chemical array, e.g., nucleic acid array, layout are provided.

30 Aspects of the invention include systems having: (a) a communication module having: (i) an input manager for receiving array request information from a user; and (ii) an output manager for providing an array layout to a user; and (b) a processing module, where the processing module has an array layout

developer configured to develop an array layout based on information that includes array request information received from a user. A feature of aspects of the invention is that the processing module is further configured to provide at least one of the following additional features. In certain embodiments, an  
5 additional feature that is provided is a collaboration manager, where the collaboration manager is a functionality configured to allow at least two different users to jointly provide array request information to said array layout developer. In certain embodiments, an additional feature that is provided is a security manager configured to control information transfer in a predetermined manner  
10 between at least two different users via said system. In certain embodiments, an additional feature that is provided is a vendor manager configured to provide access by a user to a service provided by at least one vendor. In certain of these embodiments, the service comprises providing sequences of probes for one or more targets.

15 In certain embodiments, the communication module provides for remote communication between a user and the processor module, e.g., via the Internet, where a graphical user interface (GUI) between a user and a processing module may be provided.

Aspects of the invention further include methods of obtaining an array  
20 layout for a chemical array. These methods include inputting array request information into a system (e.g., via the internet and/or a GUI) of the present invention and obtaining an array layout from the system.

Aspects of the invention also include methods of producing an array layout for a chemical array, e.g., via the internet and/or through a GUI, into a system  
25 according to the invention; and obtaining an array layout from said system to produce said array layout.

In certain embodiments, the methods further include ordering a chemical array having said array layout via said system from a chemical array vendor. In certain embodiments, the methods include receiving said chemical array from  
30 said chemical array vendor. In one aspect, users can share probe group information to order individual or shared arrays. In another aspect, users can exchange information relating to probe groups and/or array layout design. In one aspect, a customer user of the system can select probe group(s) from two or

more probe vendors and design a single array comprising the probe groups. In another aspect, the customer user may design a plurality of arrays using probe groups from the same or different vendor users of the system.

Also provided are computer program products comprising a computer readable storage medium having a computer program stored thereon, wherein said computer program, when loaded onto a computer, operates said computer to:

- (a) receive chemical array request information from a user;
- (b) develop a chemical array layout based on said received information;

10 and

(c) provide said array layout to said user. A feature of embodiments of the invention is that the computer program further controls the computer to perform at least one of the following additional tasks:

(i) allow at least two different users to jointly provide said chemical array request information;

15

(ii) control system mediated information communication in a predetermined manner between at least two different users; and

(iii) provide access by a user to a service provided by at least one vendor.

20

In one embodiment, the invention provides an online service that provides users with the ability to create chemical array layouts (e.g., such as DNA array layouts) and run search queries against a database of probe sequences, probe groups, and/or chemical arrays. In one aspect, the invention provides a system that allows users to search for desired results, save the results, compare and contrast different search results, customize their own array designs, download data, and order stock or custom slides directly from a vendor user of the system in a own password-protected area.

25

## BRIEF DESCRIPTIONS OF THE DRAWINGS

FIG. 1 illustrates a substrate carrying multiple arrays, such as may be fabricated by methods of the present invention;

30

FIG. 2 is an enlarged view of a portion of FIG. 1 showing multiple ideal spots or features;

FIG. 3 is an enlarged illustration of a portion of the substrate in FIG. 2;

FIG. 4 provides a functional block diagram of a session manager according to an embodiment of the present invention;

FIG. 5 schematically illustrates some methods and apparatus of the present invention;

FIG. 6 is a schematic diagram illustrating a collaboration manager functionality of the present invention;

FIG. 7 is a schematic diagram illustrating a vendor manager functionality of the present invention; and

FIG. 8 is a schematic diagram illustrating a fabrication station of the present invention.

## DEFINITIONS

Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Still, certain elements are defined below for the sake of clarity and ease of reference.

By “array layout” is meant a collection of information, e.g., in the form of a file, which represents the location of probes that have been assigned to specific features of an array format.

The phrase “array format” refers to a format that defines an array by feature number, feature size, cartesian coordinates of each feature, and distance that exists between features within a given array.

The phrase “array request information” is use broadly to encompass any type of information/data that is employed in developing an array layout, where representative types of array request information include, but are not limited to: probe content identifiers, e.g., in the form of probe sequence, gene name, accession number, annotation, etc.; array function information, e.g., in the form of types of genes to be studied using the array, such as genes from a specific species (e.g., mouse, human), genes associated with specific tissues (e.g., liver, brain, cardiac), genes associated with specific physiological functions, (e.g., apoptosis, stress response), genes associated with disease states (e.g., cancer, cardiovascular disease), etc.; array format information, e.g., feature number,

feature size, cartesian coordinates of each feature, and distance that exists between features within a given array; etc.

A "data element" represents a property of a probe sequence, which can include the base composition of the probe sequence. Data elements can also include representations of other properties of probe sequences, such as expression levels in one or more tissues, interactions between a sequence (and/or its encoded products), and other molecules, a representation of copy number, a representation of the relationship between its activity (or lack thereof) in a cellular pathway (e.g., a signaling pathway) and a physiological response, sequence similarity to other probe sequences, a representation of its function, a representation of its modified, processed, and/or variant forms, a representation of splice variants, the locations of introns and exons, functional domains etc. A data element can be represented for example, by an alphanumeric string (e.g., representing bases), by a number, by "plus" and "minus" symbols or other symbols, by a color hue, by a word, or by another form (descriptive or nondescriptive) suitable for computation, analysis and/or processing for example, by a computer or other machine or system capable of data integration and analysis.

As used herein, the term "data structure" is intended to mean an organization of information, such as a physical or logical relationship among data elements, designed to support specific data manipulation functions, such as an algorithm. The term can include, for example, a list or other collection type of data elements that can be added, subtracted, combined or otherwise manipulated. Exemplarily, types of data structures include a list, linked-list, doubly linked-list, indexed list, table, matrix, queue, stack, heap, dictionary, flat file databases, relational databases, local databases, distributed databases, thin client databases and tree. The term also can include organizational structures of information that relate or correlate, for example, data elements from a plurality of data structures or other forms of data management structures. A specific example of information organized by a data structure of the invention is the association of a plurality of data elements relating to a gene, e.g., its sequence, expression level in one or more tissues, copy number, activity states (e.g., active or non-active in one or more tissues), its modified, processed and/or variant forms, splice variants

encoded by the gene, the locations of introns and exons, functional domains, interactions with other molecules, function, sequence similarity to other probe sequences, etc. A data structure can be a recorded form of information (such as a list) or can contain additional information (e.g., annotations) regarding the information contained therein. A data structure can include pointers or links to resources external to the data structure (e.g., such as external databases). In one aspect, a data structure is embodied in a tangible form, e.g. is stored or represented in a tangible medium (such as a computer readable medium).

The term "object" refers to a unique concrete instance of an abstract data type, a class (that is, a conceptual structure including both data and the methods to access it) whose identity is separate from that of other objects, although it can "communicate" with them via messages. In some occasions, some objects can be conceived of as a subprogram which can communicate with others by receiving or giving instructions based on its, or the others' data or methods. Data can consist of numbers, literal strings, variables, references, etc. In addition to data, an object can include methods for manipulating data. In certain instances, an object may be viewed as a region of storage. In the present invention, an object typically includes a plurality of data elements and methods for manipulating such data elements.

A "relation" or "relationship" is an interaction between multiple data elements and/or data structures and/or objects. A list of properties may be attached to a relation. Such properties may include name, type, location, etc. A relation may be expressed as a link in a network diagram. Each data element may play a specific "role" in a relation.

As used herein, an "annotation" is a comment, explanation, note, link, or metadata about a data element, data structure or object, or a collection thereof. Annotations may include pointers to external objects or external data. An annotation may optionally include information about an author who created or modified the annotation, as well as information about when that creation or modification occurred. In one embodiment, a memory comprising a plurality of data structures organized by annotation category provides a database through which information from multiple databases, public or private, may be accessed, assembled, and processed. Annotation tools include, but are not limited to,

software such as BioFerret (available from Agilent Technologies, Inc., Palo Alto, California), which is described in detail in Application Serial No. 10/033,823 filed December 19, 2001 and titled "Domain-Specific Knowledge-Based Metasearch System and Methods of Using." Such tools may be used to generate a list of  
5 associations between genes from scientific literature and patent publications.

As used herein an "annotation category" is a human readable string to annotate the logical type the object comprising its plurality of data elements represents. Data structures that contain the same types and instances of data elements may be assigned identical annotations, while data structures that  
10 contain different types and instances of data elements may be assigned different annotations.

Different annotation categories can be generated as needed by a user.

As used herein, a "collaboration" refers to a process by which one or more individuals work together to generate an output such as to generate a probe  
15 group or array layout. A plurality of users may contribute (e.g., create, modify, add to, remove from) a collaboration space (e.g., a set of objects in the database) in the system.

As used herein, "sharing" refers to providing access to one or more users to the memory of the system. For example, the results of a collaboration or the  
20 output of an individual may be shared. In certain aspects, users who share in a collaboration space are not collaborators.

As used herein, a "consortium" comprises a group of users (e.g., individuals or organizations) in a collaboration or group of users sharing access to a design. In one aspect, the system according to the invention enables  
25 consortia members across organizations to work together. In certain aspects of the invention, users can exchange information (e.g., data relating to probe groups and sequence(s) contained therein, array layouts, and the like).

As used herein, a "probe sequence identifier" or an "identifier corresponding to a probe sequence" refers to a string of one or more characters  
30 (e.g., alphanumeric characters), symbols, images or other graphical representation(s) associated with a probe sequence comprising a probe sequence such that the identifier provides a "shorthand" designation for the sequence. In one aspect, an identifier comprises an accession number or a

clone number. An identifier may comprise descriptive information. For example, an identifier may include a reference citation or a portion thereof.

As used herein "probe request information" refers to any type of information that is employed to obtain one or more probes, and may comprise one or more search terms, key words, accession numbers, or probe sequences. Probe request information may take a number of different forms, such as sequence information, location identifier information, art accepted identifier, e.g., accession no, information, etc. Likewise, probe content information may take a number of different forms, such as sequence information, location identifier information, art accepted identifier, e.g., accession no, information, etc. In one aspect, "probe content information" includes a probe sequence or an identifier associated therewith, structural, and functional genomic and/or proteomic information with respect to the probe sequence and/or identifier. In another aspect, probe content information is relevant links to reagents or kits that might be used to obtain additional probe content information (e.g., such as links to sources of primers, antibodies, binding partners, and host cells, including transgenic animals expressing the sequences or modified forms thereof, and the like). In other aspects, probe content information may include, but is not limited to information regarding cell(s) or tissue(s) in which a probe sequence is expressed and/or levels of expression, information concerning physiological responses of a cell or tissue in which the sequence is expressed (e.g., whether the cell or tissue is from a patient with a disease), chromosomal location information, copy number information, information relating to similar sequences (e.g., homologous, paralogous or orthologous sequences). Additional probe content information can include frequency of the sequence in a population, information relating to polymorphic variants of the probe sequence (e.g., such as SNPs), information relating to splice variants (e.g., tissues, individuals in which such variants are expressed), and or demographic information relating to individual(s) in which the sequence is found.

The phrase "best-fit" refers to a resource allocation scheme that determines the best result in response to input data. The definition of 'best' may vary depending on a given set of predetermined parameters, such as sequence identity limits, signal intensity limits, cross-hybridization limits,  $T_m$ , base

composition limits, probe length limits, distribution of bases along the length of the probe, distribution of nucleation points along the length of the probe (e.g., regions of the probe likely to participate in hybridization, secondary structure parameters, etc. In one aspect, the system considers predefined thresholds. In  
5 another aspect, the system rank-orders fit. In a further aspect, the user defines his or her own thresholds, which may or may not include system-defined threshold.

The term "biomolecule" means any organic or biochemical molecule, group or species of interest that may be formed in an array on a substrate surface.  
10 Exemplary biomolecules include peptides, proteins, amino acids and nucleic acids.

The term "peptide" as used herein refers to any compound produced by amide formation between a carboxyl group of one amino acid and an amino group of another group.

15 The term "oligopeptide" as used herein refers to peptides with fewer than about 10 to 20 residues, i.e. amino acid monomeric units.

The term "polypeptide" as used herein refers to peptides with more than about 10 to about 20 residues. The terms "polypeptide" and "protein" may be used interchangeably.

20 The term "protein" as used herein refers to polypeptides of specific sequence of more than about 50 residue and includes D and L forms, modified forms, etc.

The term "nucleic acid" as used herein means a polymer composed of nucleotides, e.g., deoxyribonucleotides or ribonucleotides, or compounds  
25 produced synthetically (e.g., PNA as described in U.S. Patent No. 5,948,902 and the references cited therein) which can hybridize with naturally occurring nucleic acids in a sequence specific manner analogous to that of two naturally occurring nucleic acids, e.g., can participate in Watson-Crick base pairing interactions.

The terms "nucleoside" and "nucleotide" are intended to include those  
30 moieties that contain not only the known purine and pyrimidine base moieties, but also other heterocyclic base moieties that have been modified. Such modifications include methylated purines or pyrimidines, acylated purines or pyrimidines, or other heterocycles. In addition, the terms "nucleoside" and

"nucleotide" include those moieties that contain not only conventional ribose and deoxyribose sugars, but other sugars as well. Modified nucleosides or nucleotides also include modifications on the sugar moiety, e.g., wherein one or more of the hydroxyl groups are replaced with halogen atoms or aliphatic groups, or are functionalized as ethers, amines, or the like.

The terms "ribonucleic acid" and "RNA" as used herein refer to a polymer composed of ribonucleotides.

The terms "deoxyribonucleic acid" and "DNA" as used herein mean a polymer composed of deoxyribonucleotides.

The term "oligonucleotide" as used herein denotes single stranded nucleotide multimers of from about 10 to 100 nucleotides and up to 200 nucleotides in length.

A "biopolymer" is a polymer of one or more types of repeating units. Biopolymers are typically found in biological systems (although they may be made synthetically) and may include peptides or polynucleotides, as well as such compounds composed of or containing amino acid analogs or non-amino acid groups, or nucleotide analogs or non-nucleotide groups. This includes polynucleotides in which the conventional backbone has been replaced with a non-naturally occurring or synthetic backbone, and nucleic acids (or synthetic or naturally occurring analogs) in which one or more of the conventional bases has been replaced with a group (natural or synthetic) capable of participating in Watson-Crick type hydrogen bonding interactions. Polynucleotides include single or multiple stranded configurations, where one or more of the strands may or may not be completely aligned with another. For example, a "biopolymer" may include DNA (including cDNA), RNA, oligonucleotides, and PNA and other polynucleotides as described in U.S. Pat. No. 5,948,902 and references cited therein (all of which are incorporated herein by reference), regardless of the source.

A "biomonomer" references a single unit, which can be linked with the same or other biomonomers to form a biopolymer (e.g., a single amino acid or nucleotide with two linking groups, one or both of which may have removable protecting groups).

An "array," or "chemical array" used interchangeably includes any one-dimensional, two-dimensional or substantially two-dimensional (as well as a three-dimensional) arrangement of addressable regions bearing a particular chemical moiety or moieties (such as ligands, e.g., biopolymers such as polynucleotide or oligonucleotide sequences (nucleic acids), polypeptides (e.g., proteins), carbohydrates, lipids, etc.) associated with that region. In the broadest sense, the arrays of many embodiments are arrays of polymeric binding agents, where the polymeric binding agents may be any of: polypeptides, proteins, nucleic acids, polysaccharides, synthetic mimetics of such biopolymeric binding agents, etc. In many embodiments of interest, the arrays are arrays of nucleic acids, including oligonucleotides, polynucleotides, cDNAs, mRNAs, synthetic mimetics thereof, and the like. Where the arrays are arrays of nucleic acids, the nucleic acids may be covalently attached to the arrays at any point along the nucleic acid chain, but are generally attached at one of their termini (e.g. the 3' or 5' terminus). Sometimes, the arrays are arrays of polypeptides, e.g., proteins or fragments thereof.

Any given substrate may carry one, two, four or more or more arrays disposed on a front surface of the substrate. Depending upon the use, any or all of the arrays may be the same or different from one another and each may contain multiple spots or features. A typical array may contain more than ten, more than one hundred, more than one thousand more ten thousand features, or even more than one hundred thousand features, in an area of less than 20 cm<sup>2</sup> or even less than 10 cm<sup>2</sup>. For example, features may have widths (that is, diameter, for a round spot) in the range from a 10 μm to 1.0 cm. In other embodiments each feature may have a width in the range of 1.0 μm to 1.0 mm, usually 5.0 μm to 500 μm, and more usually 10 μm to 200 μm. Non-round features may have area ranges equivalent to that of circular features with the foregoing width (diameter) ranges. At least some, or all, of the features are of different compositions (for example, when any repeats of each feature composition are excluded the remaining features may account for at least 5%, 10%, or 20% of the total number of features). Interfeature areas will typically (but not essentially) be present which do not carry any polynucleotide (or other biopolymer or chemical moiety of a type of which the features are composed). Such interfeature areas

typically will be present where the arrays are formed by processes involving drop deposition of reagents but may not be present when, for example, light directed synthesis fabrication processes are used. It will be appreciated though, that the interfeature areas, when present, could be of various sizes and configurations.

5 Each array may cover an area of less than 100 cm<sup>2</sup>, or even less than 50 cm<sup>2</sup>, 10 cm<sup>2</sup> or 1 cm<sup>2</sup>. In many embodiments, the substrate carrying the one or more arrays will be shaped generally as a rectangular solid (although other shapes are possible), having a length of more than 4 mm and less than 1 m, usually more than 4 mm and less than 600 mm, more usually less than 400 mm; a width of  
10 more than 4 mm and less than 1 m, usually less than 500 mm and more usually less than 400 mm; and a thickness of more than 0.01 mm and less than 5.0 mm, usually more than 0.1 mm and less than 2 mm and more usually more than 0.2 and less than 1 mm. With arrays that are read by detecting fluorescence, the substrate may be of a material that emits low fluorescence upon illumination with  
15 the excitation light. Additionally in this situation, the substrate may be relatively transparent to reduce the absorption of the incident illuminating laser light and subsequent heating if the focused laser beam travels too slowly over a region. For example, substrate 10 may transmit at least 20%, or 50% (or even at least 70%, 90%, or 95%), of the illuminating light incident on the front as may be  
20 measured across the entire integrated spectrum of such illuminating light or alternatively at 532 nm or 633 nm.

Arrays may be fabricated using drop deposition from pulse jets of either precursor units (such as nucleotide or amino acid monomers) in the case of *in situ* fabrication, or the previously obtained biomolecule, e.g., polynucleotide.

25 Such methods are described in detail in, for example, the previously cited references including US 6,242,266, US 6,232,072, US 6,180,351, US 6,171,797, US 6,323,043, U.S. Patent Application Serial No. 09/302,898 filed April 30, 1999 by Caren et al., and the references cited therein. Other drop deposition methods can be used for fabrication, as previously described herein.

30 An exemplary chemical array is shown in Figs. 1-3, where the array shown in this representative embodiment includes a contiguous planar substrate 110 carrying an array 112 disposed on a rear surface 111b of substrate 110. It will be appreciated though, that more than one array (any of which are the same or

different) may be present on rear surface 111b, with or without spacing between such arrays. That is, any given substrate may carry one, two, four or more arrays disposed on a front surface of the substrate and depending on the use of the array, any or all of the arrays may be the same or different from one another and each may contain multiple spots or features. The one or more arrays 112 usually cover only a portion of the rear surface 111b, with regions of the rear surface 111b adjacent the opposed sides 113c, 113d and leading end 113a and trailing end 113b of slide 110, not being covered by any array 112. A front surface 111a of the slide 110 does not carry any arrays 112. Each array 112 can be designed for testing against any type of sample, whether a trial sample, reference sample, a combination of them, or a known mixture of biopolymers such as polynucleotides. Substrate 110 may be of any shape, as mentioned above. As mentioned above, array 112 contains multiple spots or features 116 of biopolymers, e.g., in the form of polynucleotides. As mentioned above, all of the features 116 may be different, or some or all could be the same. The interfeature areas 117 could be of various sizes and configurations. Each feature carries a predetermined biopolymer such as a predetermined polynucleotide (which includes the possibility of mixtures of polynucleotides). It will be understood that there may be a linker molecule (not shown) of any known types between the rear surface 111b and the first nucleotide.

Substrate 110 may carry on front surface 111a, an identification code, e.g., in the form of bar code (not shown) or the like printed on a substrate in the form of a paper label attached by adhesive or any convenient means. The identification code contains information relating to array 112, where such information may include, but is not limited to, an identification of array 112, i.e., layout information relating to the array(s), etc.

In those embodiments where an array includes two more features immobilized on the same surface of a solid support, the array may be referred to as addressable. An array is "addressable" when it has multiple regions of different moieties (e.g., different polynucleotide sequences) such that a region (i.e., a "feature" or "spot" of the array) at a particular predetermined location (i.e., an "address") on the array will detect a particular target or class of targets (although a feature may incidentally detect non-targets of that feature). Array

features are typically, but need not be, separated by intervening spaces. In the case of an array, the "target" will be referenced as a moiety in a mobile phase (typically fluid), to be detected by probes ("target probes") which are bound to the substrate at the various regions. However, either of the "target" or "probe" may  
5 be the one which is to be evaluated by the other (thus, either one could be an unknown mixture of analytes, e.g., polynucleotides, to be evaluated by binding with the other).

An array "assembly" includes a substrate and at least one chemical array, e.g., on a surface thereof. Array assemblies may include one or more chemical  
10 arrays present on a surface of a device that includes a pedestal supporting a plurality of prongs, e.g., one or more chemical arrays present on a surface of one or more prongs of such a device. An assembly may include other features (such as a housing with a chamber from which the substrate sections can be removed). "Array unit" may be used interchangeably with "array assembly".

15 The term "monomer" as used herein refers to a chemical entity that can be covalently linked to one or more other such entities to form a polymer. Of particular interest to the present application are nucleotide "monomers" that have first and second sites (e.g., 5' and 3' sites) suitable for binding to other like monomers by means of standard chemical reactions (e.g., nucleophilic  
20 substitution), and a diverse element which distinguishes a particular monomer from a different monomer of the same type (e.g., a nucleotide base, etc.). In the art synthesis of nucleic acids of this type utilizes an initial substrate-bound monomer that is generally used as a building-block in a multi-step synthesis procedure to form a complete nucleic acid.

25 The term "oligomer" is used herein to indicate a chemical entity that contains a plurality of monomers. As used herein, the terms "oligomer" and "polymer" are used interchangeably, as it is generally, although not necessarily, smaller "polymers" that are prepared using the functionalized substrates of the invention, particularly in conjunction with combinatorial chemistry techniques.  
30 Examples of oligomers and polymers include polydeoxyribonucleotides (DNA), polyribonucleotides (RNA), other polynucleotides which are C-glycosides of a purine or pyrimidine base. In the practice of the instant invention, oligomers will

generally comprise about 2-60 monomers, preferably about 10-60, more preferably about 50-60 monomers.

The terms "nucleoside" and "nucleotide" are intended to include those moieties which contain not only the known purine and pyrimidine bases, but also other heterocyclic bases that have been modified. Such modifications include methylated purines or pyrimidines, acylated purines or pyrimidines, alkylated riboses or other heterocycles. In addition, the terms "nucleoside" and "nucleotide" include those moieties that contain not only conventional ribose and deoxyribose sugars, but other sugars as well. Modified nucleosides or nucleotides also include modifications on the sugar moiety, e.g., wherein one or more of the hydroxyl groups are replaced with halogen atoms or aliphatic groups, or are functionalized as ethers, amines, or the like.

"Optional" or "optionally" means that the subsequently described circumstance may or may not occur, so that the description includes instances where the circumstance occurs and instances where it does not. For example, the phrase "optionally substituted" means that a non-hydrogen substituent may or may not be present, and, thus, the description includes structures wherein a non-hydrogen substituent is present and structures wherein a non-hydrogen substituent is not present.

"Hybridizing" and "binding", with respect to polynucleotides, are used interchangeably.

The term "substrate" as used herein refers to a surface upon which marker molecules or probes, e.g., an array, may be adhered. Glass slides are the most common substrate for biochips, although fused silica, silicon, plastic and other materials are also suitable.

When two items are "associated" with one another they are provided in such a way that it is apparent one is related to the other such as where one references the other. For example, an array identifier can be associated with an array by being on the array assembly (such as on the substrate or a housing) that carries the array or on or in a package or kit carrying the array assembly. "Stably attached" or "stably associated with" means an item's position remains substantially constant where in certain embodiments it may mean that an item's position remains substantially constant and known.

A "web" references a long continuous piece of substrate material having a length greater than a width. For example, the web length to width ratio may be at least 5/1, 10/1, 50/1, 100/1, 200/1, or 500/1, or even at least 1000/1.

5 "Flexible" with reference to a substrate or substrate web, references that the substrate can be bent 180 degrees around a roller of less than 1.25 cm in radius. The substrate can be so bent and straightened repeatedly in either direction at least 100 times without failure (for example, cracking) or plastic deformation. This bending must be within the elastic limits of the material. The foregoing test for flexibility is performed at a temperature of 20 °C.

10 "Rigid" refers to a material or structure which is not flexible, and is constructed such that a segment about 2.5 by 7.5 cm retains its shape and cannot be bent along any direction more than 60 degrees (and often not more than 40, 20, 10, or 5 degrees) without breaking.

15 The terms "hybridizing specifically to" and "specific hybridization" and "selectively hybridize to," as used herein refer to the binding, duplexing, or hybridizing of a nucleic acid molecule preferentially to a particular nucleotide sequence under stringent conditions.

20 The term "stringent assay conditions" as used herein refers to conditions that are compatible to produce binding pairs of nucleic acids, e.g., surface bound and solution phase nucleic acids, of sufficient complementarity to provide for the desired level of specificity in the assay while being less compatible to the formation of binding pairs between binding members of insufficient complementarity to provide for the desired specificity. Stringent assay conditions are the summation or combination (totality) of both hybridization and wash  
25 conditions.

30 The term "stringent assay conditions" as used herein refers to conditions that are compatible to produce binding pairs of nucleic acids, e.g., surface bound and solution phase nucleic acids, of sufficient complementarity to provide for the desired level of specificity in the assay while being less compatible to the formation of binding pairs between binding members of insufficient complementarity to provide for the desired specificity. Stringent assay conditions are the summation or combination (totality) of both hybridization and wash conditions.

"Stringent hybridization conditions" and "stringent hybridization wash conditions" in the context of nucleic acid hybridization (e.g., as in array, Southern or Northern hybridizations) are sequence dependent, and are different under different experimental parameters. Stringent hybridization conditions that can be used to identify nucleic acids within the scope of the invention can include, e.g., hybridization in a buffer comprising 50% formamide, 5×SSC, and 1% SDS at 42°C, or hybridization in a buffer comprising 5×SSC and 1% SDS at 65°C, both with a wash of 0.2×SSC and 0.1% SDS at 65°C. Exemplary stringent hybridization conditions can also include a hybridization in a buffer of 40% formamide, 1 M NaCl, and 1% SDS at 37°C, and a wash in 1×SSC at 45°C. Alternatively, hybridization to filter-bound DNA in 0.5 M NaHPO<sub>4</sub>, 7% sodium dodecyl sulfate (SDS), 1 mM EDTA at 65°C, and washing in 0.1×SSC/0.1% SDS at 68°C can be employed. Yet additional stringent hybridization conditions include hybridization at 60°C or higher and 3×SSC (450 mM sodium chloride/45 mM sodium citrate) or incubation at 42°C in a solution containing 30% formamide, 1M NaCl, 0.5% sodium sarcosine, 50 mM MES, pH 6.5. Those of ordinary skill will readily recognize that alternative but comparable hybridization and wash conditions can be utilized to provide conditions of similar stringency. In certain embodiments, the stringency of the wash conditions that set forth the conditions which determine whether a nucleic acid is specifically hybridized to a surface bound nucleic acid. Wash conditions used to identify nucleic acids may include, e.g.: a salt concentration of about 0.02 molar at pH 7 and a temperature of at least about 50°C or about 55°C to about 60°C; or, a salt concentration of about 0.15 M NaCl at 72°C for about 15 minutes; or, a salt concentration of about 0.2×SSC at a temperature of at least about 50°C or about 55°C to about 60°C for about 15 to about 20 minutes; or, the hybridization complex is washed twice with a solution with a salt concentration of about 2×SSC containing 0.1% SDS at room temperature for 15 minutes and then washed twice by 0.1×SSC containing 0.1% SDS at 68°C for 15 minutes; or, equivalent conditions. Stringent conditions for washing can also be, e.g., 0.2×SSC/0.1% SDS at 42°C.

A specific example of stringent assay conditions is rotating hybridization at 65°C in a salt based hybridization buffer with a total monovalent cation

concentration of 1.5 M (e.g., as described in U.S. Patent Application No. 09/655,482 filed on September 5, 2000, the disclosure of which is herein incorporated by reference) followed by washes of 0.5×SSC and 0.1×SSC at room temperature.

5 Stringent assay conditions are hybridization conditions that are at least as stringent as the above representative conditions, where a given set of conditions are considered to be at least as stringent if substantially no additional binding complexes that lack sufficient complementarity to provide for the desired specificity are produced in the given set of conditions as compared to the above  
10 specific conditions, where by "substantially no more" is meant less than about 5-fold more, typically less than about 3-fold more. Other stringent hybridization conditions are known in the art and may also be employed, as appropriate.

"Contacting" means to bring or put together. As such, a first item is contacted with a second item when the two items are brought or put together,  
15 e.g., by touching them to each other.

"Depositing" means to position, place an item at a location-or otherwise cause an item to be so positioned or placed at a location. Depositing includes contacting one item with another. Depositing may be manual or automatic, e.g., "depositing" an item at a location may be accomplished by automated robotic  
20 devices.

By "remote location," it is meant a location other than the location at which the array (or referenced item) is present and hybridization occurs (in the case of hybridization reactions). For example, a remote location could be another location (e.g., office, lab, etc.) in the same city, another location in a different city,  
25 another location in a different state, another location in a different country, etc. As such, when one item is indicated as being "remote" from another, what is meant is that the two items are at least in different rooms or different buildings, and may be at least one mile, ten miles, or at least one hundred miles apart.

"Communicating" information references transmitting the data representing  
30 that information as signals (e.g., electrical, optical, radio signals, etc.) over a suitable communication channel (e.g., a private or public network).

"Forwarding" an item refers to any means of getting that item from one location to the next, whether by physically transporting that item or otherwise

(where that is possible) and includes, at least in the case of data, physically transporting a medium carrying the data or communicating the data.

An array "package" may be the array plus only a substrate on which the array is deposited, although the package may include other features (such as a housing with a chamber).

A "chamber" references an enclosed volume (although a chamber may be accessible through one or more ports). It will also be appreciated that throughout the present application, that words such as "top," "upper," and "lower" are used in a relative sense only.

A "computer-based system" refers to the hardware means, software means, and data storage means used to analyze the information of the present invention. The minimum hardware of the computer-based systems of the present invention comprises a central processing unit (CPU), input means, output means, and data storage means. A skilled artisan can readily appreciate that many computer-based systems are available which are suitable for use in the present invention. The data storage means may comprise any manufacture comprising a recording of the present information as described above, or a memory access means that can access such a manufacture.

A "processor" references any hardware and/or software combination which will perform the functions required of it. For example, any processor herein may be a programmable digital microprocessor such as available in the form of an electronic controller, mainframe, server or personal computer (desktop or portable). Where the processor is programmable, suitable programming can be communicated from a remote location to the processor, or previously saved in a computer program product (such as a portable or fixed computer readable storage medium, whether magnetic, optical or solid state device based). For example, a magnetic medium or optical disk may carry the programming, and can be read by a suitable reader communicating with each processor at its corresponding station.

"Computer readable medium" as used herein refers to any storage or transmission medium that participates in providing instructions and/or data to a computer for execution and/or processing. Examples of storage media include floppy disks, magnetic tape, UBS, CD-ROM, a hard disk drive, a ROM or

integrated circuit, a magneto-optical disk, or a computer readable card such as a PCMCIA card and the like, whether or not such devices are internal or external to the computer. A file containing information may be "stored" on computer readable medium, where "storing" means recording information such that it is accessible  
5 and retrievable at a later date by a computer. A file may be stored in permanent memory.

With respect to computer readable media, "permanent memory" refers to memory that is permanently stored on a data storage medium. Permanent memory is not erased by termination of the electrical supply to a computer or  
10 processor. Computer hard-drive ROM (i.e. ROM not used as virtual memory), CD-ROM, floppy disk and DVD are all examples of permanent memory. Random Access Memory (RAM) is an example of non-permanent memory. A file in permanent memory may be editable and re-writable.

To "record" data, programming or other information on a computer  
15 readable medium refers to a process for storing information, using any such methods as known in the art. Any convenient data storage structure may be chosen, based on the means used to access the stored information. A variety of data processor programs and formats can be used for storage, e.g. word processing text file, database format, etc.

20 A "memory" or "memory unit" refers to any device which can store information for subsequent retrieval by a processor, and may include magnetic or optical devices (such as a hard disk, floppy disk, CD, or DVD), or solid state memory devices (such as volatile or non-volatile RAM). A memory or memory unit may have more than one physical memory device of the same or different  
25 types (for example, a memory may have multiple memory devices such as multiple hard drives or multiple solid state memory devices or some combination of hard drives and solid state memory devices).

Items of data are "linked" to one another in a memory when the same data input (for example, filename or directory name or search term) retrieves the linked  
30 items (in a same file or not) or an input of one or more of the linked items retrieves one or more of the others.

It will also be appreciated that throughout the present application, that words such as "cover", "base" "front", "back", "top", are used in a relative sense

only. The word "above" used to describe the substrate and/or flow cell is meant with respect to the horizontal plane of the environment, e.g., the room, in which the substrate and/or flow cell is present, e.g., the ground or floor of such a room.

5

## DETAILED DESCRIPTION OF THE INVENTION

Systems and methods for obtaining chemical array layouts are provided. The subject systems include a communications module and a processing module, where the processing module includes an array layout developer configured to develop a chemical array layout based on information that includes  
10 array request information received from a user. In certain embodiments, the processor module includes a collaboration manager configured to allow at least two different users to jointly provide chemical array request information to the array layout developer. In certain embodiments, the processor module includes a security manager configured to control system mediated information transfer in a  
15 predetermined manner between at least two different users. In certain embodiments, the processor includes a vendor manager configured to provide access by a user to a service provided by at least one vendor. In certain embodiments, the methods further include a step of fabricating an array having features arranged according to an array layout determined by the system. In  
20 certain embodiments, the methods further include shipping such arrays, e.g., to a user of the system or third party. Also provided are computer program products for executing the subject methods.

Before the present invention is described in greater detail, it is to be  
25 understood that this invention is not limited to particular embodiments described, as such may, of course, vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to be limiting, since the scope of the present invention will be limited only by the appended claims.

30 Where a range of values is provided, it is understood that each intervening value, to the tenth of the unit of the lower limit unless the context clearly dictates otherwise, between the upper and lower limit of that range and any other stated or intervening value in that stated range is encompassed within the invention.

The upper and lower limits of these smaller ranges may independently be included in the smaller ranges is also encompassed within the invention, subject to any specifically excluded limit in the stated range. Where the stated range includes one or both of the limits, ranges excluding either or both of those included limits are also included in the invention.

Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any methods and materials similar or equivalent to those described herein can also be used in the practice or testing of the present invention, the preferred methods and materials are now described.

All publications and patents cited in this specification are herein incorporated by reference as if each individual publication or patent were specifically and individually indicated to be incorporated by reference and are incorporated herein by reference to disclose and describe the methods and/or materials in connection with which the publications are cited. The citation of any publication is for its disclosure prior to the filing date and should not be construed as an admission that the present invention is not entitled to antedate such publication by virtue of prior invention. Further, the dates of publication provided may be different from the actual publication dates which may need to be independently confirmed.

It must be noted that as used herein and in the appended claims, the singular forms "a", "an", and "the" include plural referents unless the context clearly dictates otherwise. It is further noted that the claims may be drafted to exclude any optional element. As such, this statement is intended to serve as antecedent basis for use of such exclusive terminology as "solely," "only" and the like in connection with the recitation of claim elements, or use of a "negative" limitation.

As will be apparent to those of skill in the art upon reading this disclosure, each of the individual embodiments described and illustrated herein has discrete components and features which may be readily separated from or combined with the features of any of the other several embodiments without departing from the scope or spirit of the present invention. Any recited method can be carried out in the order of events recited or in any other order which is logically possible.

As summarized above, aspects of the invention include systems and methods of using the same which may be employed to produce array layouts for chemical arrays. In further describing the aspects of the invention, a review of representative system hardware/software architecture is provided, followed by a more detailed discussion of aspects of representative embodiments of the invention.

As summarized above, aspects of the invention include systems and methods for producing an array layout. Representative embodiments of the subject systems generally include the following components: (a) a communications module for facilitating information transfer between the system and one or more users, e.g., via a user computer, as described below; and (b) a processing module for performing one or more tasks in response to information received via the communications module of the system. In representative embodiments, the subject systems may be viewed as being the physical embodiment of a web portal, where the term "web portal" refers to a web site or service, e.g., as may be viewed in the form of a web page, that offers a broad array of resources and services to users via an electronic communication element, e.g., via the Internet. Each of these elements is described in greater detail below.

The subject systems may include both hardware and software components, where the hardware components may take the form of one or more platforms, e.g., in the form of servers, such that the functional elements, i.e., those elements of the system that carry out specific tasks (such as managing input and output of information, processing information, etc.) of the system may be carried out by the execution of software applications on and across the one or more computer platforms represented of the system.

The one or more platforms present in the subject systems may be any type of known computer platform or a type to be developed in the future, although they typically will be of a class of computer commonly referred to as servers. However, they may also be a main-frame computer, a work station, or other computer type. They may be connected via any known or future type of cabling or other

communication system including wireless systems, either networked or otherwise. They may be co-located or they may be physically separated. Various operating systems may be employed on any of the computer platforms, possibly depending on the type and/or make of computer platform chosen. Appropriate  
5 operating systems include Windows NT®, Sun Solaris, Linux, OS/400, Compaq Tru64 Unix, SGI IRIX, Siemens Reliant Unix, and others.

In certain embodiments, the subject devices include multiple computer platforms which may provide for certain benefits, e.g., lower costs of deployment, database switching, or changes to enterprise applications, and/or more effective  
10 firewalls. Other configurations, however, are possible. For example, as is well known to those of ordinary skill in the relevant art, so-called two-tier or N-tier architectures are possible rather than the three-tier server-side component architecture represented by, for example, E. Roman, Mastering Enterprise  
JavaBeans™ and the Java™2 Platform ( John Wiley & Sons, Inc., NY, 1999) and  
15 J. Schneider and R. Arora, Using Enterprise Java. (Que Corporation, Indianapolis, 1997).

It will be understood that many hardware and associated software or firmware components that may be implemented in a server-side architecture for Internet commerce are known and need not be reviewed in detail here.  
20 Components to implement one or more firewalls to protect data and applications, uninterruptable power supplies, LAN switches, web-server routing software, and many other components are not shown. Similarly, a variety of computer components customarily included in server-class computing platforms, as well as other types of computers, will be understood to be included but are not shown.  
25 These components include, for example, processors, memory units, input/output devices, buses, and other components noted above with respect to a user computer. Those of ordinary skill in the art will readily appreciate how these and other conventional components may be implemented.

The functional elements of system may also be implemented in  
30 accordance with a variety of software facilitators and platforms (although it is not precluded that some or all of the functions of system may also be implemented in hardware or firmware). Among the various commercial products available for implementing e-commerce web portals are BEA WebLogic from BEA Systems,

which is a so-called "middleware" application. This and other middleware applications are sometimes referred to as "application servers," but are not to be confused with application server hardware elements. The function of these middleware applications generally is to assist other software components (such as software for performing various functional elements) to share resources and coordinate activities. The goals include making it easier to write, maintain, and change the software components; to avoid data bottlenecks; and prevent or recover from system failures. Thus, these middleware applications may provide load-balancing, fail-over, and fault tolerance, all of which features will be appreciated by those of ordinary skill in the relevant art.

Other development products, such as the Java™2 platform from Sun Microsystems, Inc. may be employed in the system to provide suites of applications programming interfaces (API's) that, among other things, enhance the implementation of scalable and secure components. Various other software development approaches or architectures may be used to implement the functional elements of system and their interconnection, as will be appreciated by those of ordinary skill in the art.

In one embodiment, objects are organized within the system according to domains. Domains can include, but are not limited to, provider domains (e.g., such as vendor domains) and customer domains. The system may also include a root domain that does not belong to any particular organization (e.g., vendor or customer), but is considered as the "superuser" domain. Domains may further include sub-domains. For example, a vendor domain may include sub-domains corresponding to different product and service providers within the organization. In one aspect, a user in a higher domain may have a plurality of roles (e.g., set of privileges or permissions) and these may be applied in all lower subdomains. Generally, there will be a single superuser who belongs to the root domain who has unrestricted access to all domains and sub-domains in the system.

A user of the system may be assigned to a default domain and would be able to view all objects in the sub-domains of his/her default domain. In such a case, a user may have the same roles as defined in the default domain for all the sub-domains. A user also can have a secondary list of domains in which he/she can do work. These domains are the domains that are not part of domain

hierarchy in which the user belongs. In one aspect, each domain entry in the secondary list has a list of roles associated with it. For example, a user who is a customer of a vendor who provides arrays using the system may have a vendor domain as a secondary domain, allowing them view access to selected objects in the vendor domain which they require for ordering arrays. For example, the user may be provided access to objects associated with probe group selection and/or array layout.

In one embodiment, a user may share objects with others having access to the same domain. In another embodiment, a user may choose to provide access to those with access to sub-domains.

In another embodiment, the system provides a mechanism to restrict users from accessing a type of object category or class. For example, a user may be restricted by role to performing only those set of actions permitted by that role. Permissions may include: Read, Write, Read and Write, and Share.

In one embodiment, when a user with the required privileges wants to share an object to another domain, he/she will do so from a user interface by selecting the domain(s) to which this object needs to be shared to. In one aspect, every sharable object, has a set of domain shares which will be modified when a user changes sharing for an object. In another aspect, upon saving the object, the changes to the domain shares are automatically saved to a database table. In a further aspect, when a new user gains access to the system, all objects in at least one domain are shared with the new user with read-only privileges. For example, a customer/user may have read-only access to a vendor domain.

In one embodiment, security is defined for each object, i.e., an object may also have privileges associated with it. For example, an object, such as Probe Group 1, comprising a collection of data elements corresponding to a group of probes, may have "instance level security" identifying which users are able to perform which actions on that particular probe group. For example, a user may be assigned a role that allows them to read and write to the object Probe Group 1. In such a case, the user would be able to read output relating to the contents of Probe Group 1 but would not be able to create, delete or otherwise modify the probe group. In certain aspects, users' role(s) are saved in the system memory and can be retrieved by the system.

The system may include a mechanism to enable or disable privileges by using a configurable parameter and in one aspect includes a database table which comprises an object ID (a unique database ID for an object), a user ID (a database ID of a person who has privileges relating to the object) and a privilege ID (the privilege assigned to a user for the particular object

In another embodiment, the system includes a mechanism for maintaining an audit trail of interactions with the system by providing an auditable object. In one aspect, an auditable object has a persistent database table to store audit entries. Audit entries may be associated with a unique database ID and an object ID corresponding to the object for which the audit entry is created. The table may also include the category or class to which the object for which the audit entry is being created belongs. In certain aspects, the database may further include the some description associated with the audit entry (e.g., such as reason for the audit entry, audit date, etc). Additionally, a user ID uniquely identifying a user who performed any modification (e.g., creation, updating, deleting, etc) may be included.

In a further embodiment, the system keeps track of information relating to downloads by a user. For example, the system may identify download date, the user id, file name, and the object(s) associated with any downloaded files and store and/or output the information in the system memory. The stored information and/or output may be assigned a unique database ID and be associated with a log ID identifying the date on which the information was stored and/or output was generated.

In one embodiment, a user enters into a session with the system. A session represents a series of requests from a particular user to a particular application of the system over a certain period of time. In one aspect, the system maintains a memory of a session object's state(s). The system may rely on this information in processing a new request.

In another embodiment, the system comprises a mechanism by which an administrator of the system can monitor the number of users connected to the system at a particular time. In one aspect, an administrator can invalidate the session of any user at any time, so that the user would not be able to access the system.

A variety of interfaces may be used to implement the functions of the system. In one embodiment, in the case of web applications, a servlet container uses an HTTP Session interface to create a session between an HTTP client and an HTTP server. The session may persist for a specified time period, across  
5 more than one connection or page request from a user. In one aspect, one user may be involved in a session, and the user may visit the web application many times. However, multiple users also may be involved in a session. The server can maintain a session in many ways, such as by using cookies or rewriting URLs.

10 In another embodiment, the system comprises a session manager. The session manager acts as a factory class that may be used to generate objects, and in one aspect, related objects when a user interacts with the system. In another embodiment, information relating to all user sessions is maintained in a collection within the session manager. In a further embodiment, one session  
15 manager instance is associated with one application in the system. In still a further embodiment, session instances are associated with session manager instances. This structure ensures that there are collections of instances per application in the system.

The session manager may have one or more of the following properties.  
20 The session manager may comprise a collection of all Session objects for all current users using the system or an application of the system. In one aspect, the collection is in the form of a Hashtable.

In one embodiment, the system contains a plurality of different application objects. Application objects comprise object representations of underlying  
25 database tables. In one aspect, each application has a context associated with it. Context is a logical area of the application, which contains the configuration information for the application. This information can be accessed within that application via this context.

For example, in one embodiment, the system comprises an application  
30 bootstrap framework, which comprises a set of classes and a configuration file. In one aspect, the configuration file contains configuration information for each application. The application bootstrapping mechanism starts working when the system starts up for the first time. When system starts up, a system initialization

program (e.g., start up Servlet) instantiates an instance of Application object per application in the system. The first request to the application server will check whether application context for the named application is there or not. If application context is not present then it creates one. In one aspect, the application bootstrap framework communicates with an object/relationship mapping means in the system, assisting a user to identify object categories associated with a user query. In another aspect, in response to the identification of object categories, an output (e.g., such as a display on a graphical user interface) is generated

10 In one embodiment, the system includes an event generation and processing framework. Whenever an action takes place on an object in the system, the system generates an event. The object that generates this event is called as the event source. In one aspect, when events occur, a user with requisite permissions is notified for these events. In certain aspects, to get an event notification, the user must register him/herself for that type of event. The user will get notifications only for those types of events for which the user has registered. For this, the system maintains a queue of the events, which contains only those events for which at least one user has registered. This queue is then processed periodically and notifications are sent to the users, e.g., by email. In one embodiment, the event notification framework generates events and adds them to the event queue, while the event processing framework processes the events from the event queue and then sends the notifications. See Figure 4.

25 In one aspect, events supported by system application(s) are pre-configured. For example, the system memory can include a database of all supported (e.g., pre-configured events). In one aspect, the database includes a table comprising an event ID uniquely identifying a supported event (e.g., an annotation update), an action name for the event (e.g., "Annotation Update"), and name of an action that will be executed during post-processing of an event. The table may be a hashtable collection which may be associated with a particular user session by a session ID. In one aspect, the event manager allows a user to create, add and/or notify a user about events.

The Event Manager may include a mechanism for providing an output to a user which may include, but is not limited to the name of the event, an ID for an

event uniquely identifying the event in the database, date of the event, content of an message to the user describing the event, type of event (e.g., triggered or periodic), and the like.

5 In certain aspects, a user may have an event manager associated with that particular user's events.

In a further aspect, the system comprises a Hashtable collection which contains a key-value pair of application name and session manager instance associated with an application. This collection is useful for identifying session manager instances for all applications in the system.

10 In one embodiment, a system according to the invention creates a session manager for an application if one did not already exist. In one aspect, the system may output data relating to all the session manager instances that are associated with the system (e.g., for all applications of the system). Similarly, the system may output information relating to the collection of session instances associated  
15 with any given session manager. The system may further remove a session from a session collection as well as invalidate a user session.

In one embodiment, the system further includes an instructional module that executes instructions from a computer program product for displaying Web pages that instruct a user how to use and interact with the system to order probe  
20 groups and/or arrays and/or associated reagents. In one aspect, the instructional module provides a tutorial page, explaining the purpose of the module (e.g., to provide instructions for designing and/or ordering arrays, and/or optionally, defining terms (e.g., probe groups, arrays, array layouts, annotations). Additional Web pages or sections of web pages can be provided to describe and provide  
25 examples of various system functions (e.g., such as searching, uploading probes, downloading probes, etc.) and can provide interactive sessions to illustrate system functions. Such sessions can include displaying information relating to searching for information about probes, identifying probes, uploading probes, downloading probes, demonstrating sorting, viewing, saving search results,  
30 providing tutorials for generating an array layout, and the like. The instructional module can include a variety of graphics, including text, images, animation and can also provide accompanying voiceovers.

Figure 6 provides a view of a representative system according to an embodiment of the subject invention. In Figure 5, system 500 includes communications module 520 and processing module 530, where each module may be present on the same or different platforms, e.g., servers, as described  
5 above. The communications module includes the input manager 522 and output manager 524 functional elements.

Input manager 522 receives information, e.g., request information, from a user e.g., over the Internet. Input manager 522 processes and forwards this information to the processing module 530. These functions are performed in  
10 accordance with known techniques common to the operation of Internet servers, also commonly referred to in similar contexts as presentation servers. Another of the functional elements of communications module 520 is output manager 524. Output manager 524 provides information assembled by processing module, e.g., array layout and/or probe related content, to a user, e.g., over the Internet, also in  
15 accordance with those known techniques. The presentation of data by the output manager may be implemented in accordance with a variety of known techniques. As some examples, data may include HTML or XML documents, email or other files, or data in other forms. The data may include Internet URL addresses so that a user may retrieve additional HTML, XML, or other documents or data from  
20 remote sources.

The communications module 520 may be operatively connected to a user computer 510, which provides a vehicle for a user to interact with the system 500. User computer 510, shown in Figure 6, may be a computing device specially designed and configured to support and execute any of a multitude of different  
25 applications. Computer 510 also may be any of a variety of types of general-purpose computers such as a personal computer, network server, workstation, or other computer platform now or later developed. Computer 510 typically includes known components such as a processor, an operating system, a graphical user interface (GUI) controller, a system memory, memory storage devices, and input-  
30 output controllers. It will be understood by those skilled in the relevant art that there are many possible configurations of the components of computer 510 and that some components are not listed above, such as cache memory, a data backup unit, and many other devices. The processor may be a commercially

available processor such as a Pentium® processor made by Intel Corporation, a SPARC® processor made by Sun Microsystems, or it may be one of other processors that are or will become available. The processor executes the operating system, which may be, for example, a Windows®-type operating system (such as Windows NT®4.0 with SP6a) from the Microsoft Corporation; a Unix® or Linux-type operating system available from many vendors; another or a future operating system; or some combination thereof. The operating system interfaces with firmware and hardware in a well-known manner, and facilitates the processor in coordinating and executing the functions of various computer programs that may be written in a variety of programming languages, such as Java, Perl, C++, other high level or low level languages, as well as combinations thereof, as is known in the art. The operating system, typically in cooperation with the processor, coordinates and executes functions of the other components of the computer. The operating system also provides scheduling, input-output control, file and data management, memory management, and communication control and related services, all in accordance with known techniques.

The system memory may be any of a variety of known or future memory storage devices. Examples include any commonly available random access memory (RAM), magnetic medium such as a resident hard disk or tape, an optical medium such as a read and write compact disc, or other memory storage device. The memory storage device may be any of a variety of known or future devices, including a compact disk drive, a tape drive, a removable hard disk drive, or a diskette drive. Such types of memory storage devices typically read from, and/or write to, a program storage medium (not shown) such as, respectively, a compact disk, magnetic tape, removable hard disk, or floppy diskette. Any of these program storage media, or others now in use or that may later be developed, may be considered a computer program product. As will be appreciated, these program storage media typically store a computer software program and/or data. Computer software programs, also called computer control logic, typically are stored in system memory and/or the program storage device used in conjunction with the memory storage device.

In some embodiments, a computer program product is described comprising a computer usable medium having control logic (computer software

program, including program code) stored therein. The control logic, when executed by the processor the computer, causes the processor to perform functions described herein. In other embodiments, some functions are implemented primarily in hardware using, for example, a hardware state machine.

- 5 Implementation of the hardware state machine so as to perform the functions described herein will be apparent to those skilled in the relevant arts.

The input-output controllers of the computer could include any of a variety of known devices for accepting and processing information from a user, whether a human or a machine, whether local or remote. Such devices may include, for example, modem cards, network interface cards, sound cards, or other types of controllers for any of a variety of known input devices. Output controllers of input-output controllers could include controllers for any of a variety of known display devices for presenting information to a user, whether a human or a machine, whether local or remote. If one of the display devices provides visual information, this information typically may be logically and/or physically organized as an array of picture elements, sometimes referred to as pixels. A graphical user interface (GUI) controller may comprise any of a variety of known or future software programs for providing graphical input and output interfaces between the computer 510 and a user, and for processing user inputs. The functional elements of the computer 510 may communicate with each other via system bus. Some of these communications may be accomplished in alternative embodiments using network or other types of remote communications.

During use, a user employs the user computer to enter information into and retrieve information from the system. As shown in Figure 5, Computer 510 is coupled via network cable 400 to the system 500. Additional computers of other users in a local or wide-area network including an Intranet, the Internet, or any other network may also be coupled to system 500 via cable 400. It will be understood that cable 400 is merely representative of any type of network connectivity, which may involve cables, transmitters, relay stations, network servers, and many other components not shown but evident to those of ordinary skill in the relevant art. Via user computer 510, a user may operate a web browser served by a user-side Internet client to communicate via Internet with

system 500. System 500 may similarly be in communication over Internet with other users and/or networks of users, as desired.

As reviewed above, the systems include various functional elements that carry out specific tasks on the platforms in response to information introduced  
5 into the system by one or more users. In Figure 5, elements 532, 534 and 536 represent three different functional elements of processing module 530. Representative functional elements that may be carried out by the processing module are now reviewed in greater detail below.

The subject processing modules typically include at least one functional  
10 element that generates an array layout based on information received from one or more users. This functional element is conveniently referred to herein as an array layout developer. The array layout developer is configured to develop a chemical array layout in response to information received from one or more users, where the information received from the one or more users typically  
15 includes array request information. As reviewed above, by "array layout" is meant a collection of information, e.g., in the form of a file, that represents the location of probes that have been assigned to specific features of an array format. The phrase "array format" refers to a format that defines an array by feature number, feature size, cartesian coordinates of each feature, and distance that exists  
20 between features within a given array. The phrase "array request information" is used broadly to encompass any type of information/data that is employed in developing an array layout, where representative types of array request information include, but are not limited to: probe content identifiers, e.g., in the form of probe sequence, gene name, accession number, annotation, etc.; array  
25 function information, e.g., in the form of types of genes to be studied using the array, such as genes from a specific species (e.g., mouse, human), genes associated with specific tissues (e.g., liver, brain, cardiac), genes associated with specific physiological functions, (e.g., apoptosis, stress response), genes associated with disease states (e.g., cancer, cardiovascular disease), etc.; array  
30 format information, e.g., feature number, feature size, cartesian coordinates of each feature, and distance that exists between features within a given array; etc. As such, the array layout developer of the processing modules of the subject

systems is a functional element that produces an array layout in response to receiving array request information.

In certain embodiments, the processing module is configured to include a collaboration manager. The collaboration manager is a functional element that is configured to allow at least two different users to jointly provide array request information to the array layout developer. By two or more is meant a plurality of at least two, where the number may be 3 or more, 5 or more, 10 or more, etc., up to "n" different users, where "n" may be as higher as 25 or more, e.g., 50 or more, 100 or more, etc. In certain embodiments, the collaboration manager provides equal ability (i.e., access) for all users to participate in the generation of the array layout, and specifically in the inputting of request information to the array layout developer. In these embodiments, each user will have an undifferentiated, e.g., in the form of unrestricted, ability to introduce array request information to the array layout developer. In yet other embodiments, the collaboration manager differentiates users into one or more categories that have differing abilities to provide array request information to the array layout developer. For example, the collaboration manager may control the input from a first user such that the first user is only able to provide array request information which can be viewed as array layout content/format information, such that the first user may be viewed as the array layout creator. In such embodiments, the first user may input into the system feature content information, array format information, etc., but cannot participate in other stages of the array layout development process, such as reviewing of an array layout and finalization of an array layout. In such embodiments, the collaboration manager may control input from a second user such that the second user is only able to review candidate array layouts produced in response to array request information provided by a first user. The second user cannot create array layouts, but only review and possibly edit array layouts that are made available to the user by the system. In such embodiments, the second user may review array layouts to determine whether they meet one or more predetermined criteria, such as content, cost, format, etc. The collaboration manager may control input from a third user such that the third user is only able to finalize and accept a given array layout. As such, the third user is not able to provide array request information that is used to produce array layouts, but

instead is only able to approve or disapprove final array layouts, e.g., ones that have been reviewed by the second user. The above embodiments of the collaboration manager are merely representative, such that the collaboration manager may control input of various users in a multitude of different ways, as  
5 desired.

The Collaboration Manager enables Users to insert, remove, modify, and synthesize data in a "Collaboration Space". The Collaboration Space is an environment in which Users can collectively create a Probe Group or Array Layout by versioning these elements with the addition and subtraction of content.  
10 The Collaboration Manager interprets and supports the business rules (permissions, active functions) of the User Organization (Company/Institution), as well as the individual User, all within the Collaboration Space. The Collaboration Space has its own business rule set that enables Users to view and edit others content, and to synthesize new content (Probes, ProbeGroups, Array Layouts,  
15 and other types of content), within this environment. Figure 6 provides a diagram of a representative Collaboration Manager.

In certain embodiments, the processing module is configured to include a security manager configured to control information transfer in a predetermined manner between at least two different users via the system. In other words, the  
20 security manager controls information exchange between two or more users having access to the system. For example, a first user may choose to store an array layout, probe group, probes, or array request information (e.g., probe information, annotation information, array design data, etc.) on the system, and provide access to such information to other users of the system, thereby  
25 "publishing" the array layout, probe group, probes, or array request information on the system. In certain embodiments, the user may choose to publish the array layout or array request information to only a subpopulation of the users of the system. A first user or system administrator can provide for a universal rule regarding access, or an individual rule at the time the information is published. In  
30 such embodiments, the security manager may be configured to provide access privileges to certain users but not others. Other users may choose, prior to using the system, that they do not wish to have access to array layouts of other users. The security manager may limit access of such users to prevent such users from

seeing array layouts of others that may be stored on the system. The above-described embodiments are merely representative of various different ways the security manager may control information exchange between two or more users of the system.

5           In certain embodiments, the processing module is configured to include a vendor manager. The vendor manager is a functional element that is configured to provide access by a user to a service provided by at least one vendor, e.g., of array content. As such, the vendor manager may provide access by a user to information posted on the system by a vendor, where the information may take  
10   the form of array content (such as probe sequences), an offer to perform a service, e.g., generate an array according to a given layout, a link to a vendor website, etc. In other words, the vendor manager enables transfer of information between one or more vendors and a user. As such, the vendor manager provides for communication between a user and one or more vendors. The vendor may be  
15   the same or different than the operator of the system, and information from multiple vendors may be controlled by the vendor manager, e.g., 2 or more, 5 or more, 10 or more, 20 or more, etc. In order to be displayed by the vendor manager to a user, the vendors may pay a fee to the system operator. As an example, the processing module enables Users of the System to obtain  
20   commercially available probes, target sequence data, or related annotation for inclusion in the Array layouts that they create. The processing module, in certain embodiments, manages business rules between the vendor and User(s) of the System, including, but not limited to, payment for data access, payment on a per data record download, and other vendor data access rules, conditions,  
25   regulations, and restrictions. Figure 7 provides a functional diagram of a representative vendor manager.

As summarized above, the systems of the invention receive array request information from a user and generate an array layout therefrom. The generated array layout is, in representative embodiments, forwarded to the user for  
30   evaluation and use. As such, the systems find use in at least generating array layouts. The array layouts generated by the subject systems can be layouts for any type of chemical array, where in representative embodiments the array layouts are layouts for biopolymeric arrays, such as nucleic acid and amino acid

arrays. In representative embodiments, the layouts generated by the subject systems are for nucleic acid arrays.

In using the subject systems, as summarized above, a user or users input array request information into the system, e.g., via a user computer, as reviewed  
5 above. As reviewed above, the array request information may take a number of different forms, such as content information, format information, array layout review information, array layout approval information, etc. The system then takes the provided request information and ultimately generates a final array layout. The final array layout is then forwarded to the user, e.g., via the user computer. In  
10 certain embodiments, the final array layout, and even request information used to generate the same, is stored on the system in a suitable memory element, where access to the stored information may be free to other users, or controlled in some way, as managed by a security manager, described above.

In certain embodiments, the systems include an array layout functionality,  
15 as described in copending application serial no. \_\_\_\_\_ (attorney docket number 10041581-1) titled " Systems and Methods for Producing Chemical Array Layouts," and filed on even date herewith. In certain of these embodiments, the system includes an array layout developer, where the array layout developer includes a memory having a plurality of rules relating to array layout design and is  
20 configured to develop an array layout based on the application of one or more of the rules to information that includes array request information received from a user.

In certain embodiments, the systems include probe design functionality, as described in copending application serial no. \_\_\_\_\_ (attorney docket  
25 number 10040938-1) titled " Systems and Methods for Array Probe Design," and filed on even date herewith. In certain of these embodiments, the system includes a processing module configured to identify a probe sequence based on information regarding attributes of the plurality of data structures and then provide the identified probe sequence, or an identifier thereof, to a user.

30 In certain embodiments, the output manager further provides a user with information regarding how to purchase an according to the provided array layout. In certain embodiments, the information is provided in the form of an email. In certain embodiments, the information is provided in the form of web page content

on a graphical user interface in communication with the output manager. In certain embodiments, the web page content provides a user with an option to select for purchase one or more synthesized arrays. In certain embodiments, the web page content includes fields for inputting customer information. In certain  
5       embodiments, the system can store the customer information in the memory. In certain embodiments, the customer information includes one or more purchase order numbers. In certain embodiments, the customer information includes one or more purchase order numbers and the system prompts a user to select a purchase order number prior to purchasing the one or more synthesized arrays.

10           In certain embodiments, the user may choose to obtain an array having the generated array layout. Array fabrication according to an array layout can be accomplished in a number of different ways. With respect nucleic acid arrays in which the immobilized nucleic acids are covalently attached to the substrate surface, such arrays may be synthesized via in situ synthesis in which the nucleic  
15       acid ligand is grown on the surface of the substrate in a step-wise fashion and via deposition of the full ligand, e.g., in which a presynthesized nucleic acid/polypeptide, cDNA fragment, etc., onto the surface of the array.

Where the in situ synthesis approach is employed, conventional phosphoramidite synthesis protocols are typically used. In phosphoramidite  
20       synthesis protocols, the 3'-hydroxyl group of an initial 5'-protected nucleoside is first covalently attached to the polymer support, e.g., a planar substrate surface. Synthesis of the nucleic acid then proceeds by deprotection of the 5'-hydroxyl group of the attached nucleoside, followed by coupling of an incoming nucleoside-3'-phosphoramidite to the deprotected 5' hydroxyl group (5'-OH). The  
25       resulting phosphite triester is finally oxidized to a phosphotriester to complete the internucleotide bond. The steps of deprotection, coupling and oxidation are repeated until a nucleic acid of the desired length and sequence is obtained. Optionally, a capping reaction may be used after the coupling and/or after the oxidation to inactivate the growing DNA chains that failed in the previous coupling  
30       step, thereby avoiding the synthesis of inaccurate sequences.

In the synthesis of nucleic acids on the surface of a substrate, reactive deoxynucleoside phosphoramidites are successively applied, in molecular amounts exceeding the molecular amounts of target hydroxyl groups of the

substrate or growing oligonucleotide polymers, to specific cells of the high-density array, where they chemically bond to the target hydroxyl groups. Then, unreacted deoxynucleoside phosphoramidites from multiple cells of the high-density array are washed away, oxidation of the phosphite bonds joining the newly added  
5 deoxynucleosides to the growing oligonucleotide polymers to form phosphate bonds is carried out, and unreacted hydroxyl groups of the substrate or growing oligonucleotide polymers are chemically capped to prevent them from reacting with subsequently applied deoxynucleoside phosphoramidites. Optionally, the capping reaction may be done prior to oxidation.

10 With respect to actual array fabrication, in certain embodiments, the user may itself produce an array having the generated array layout. In yet other embodiments, the user may forward the array layout to a specialized array fabricator or vendor, which vendor will then fabricate the array according to the array layout.

15 In yet other embodiments, the system may be in communication with an array fabrication station, e.g., where the system operator is also an array vendor, such that the user may order an array directly through the system. In response to receiving an order from the user, the system will forward the array layout to a fabrication station, and the fabrication station will fabricate the array according to  
20 the forwarded array layout.

Arrays can be fabricated using drop deposition from pulsejets of either polynucleotide precursor units (such as monomers) in the case of *in situ* fabrication, or the previously obtained polynucleotide. Such methods are described in detail in, for example, the previously cited references including US  
25 6,242,266, US 6,232,072, US 6,180,351, US 6,171,797, US 6,323,043, U.S. Patent Application Serial No. 09/302,898 filed April 30, 1999 by Caren et al., and the references cited therein. Other drop deposition methods can be used for fabrication, as previously described herein. Also, instead of drop deposition methods, light directed fabrication methods may be used, as are known in the art.  
30 Interfeature areas need not be present particularly when the arrays are made by light directed synthesis protocols.

A representative array fabrication device and system is depicted in Figure 8. The apparatus shown includes a substrate station 120 on which can be

mounted a substrate 10. Pins or similar means (not shown) can be provided on substrate station 120 by which to approximately align substrate 10 to a nominal position thereon (with alignment marks 18 on substrate 10 being used for more refined alignment). Substrate station 120 can include a vacuum chuck connected to a suitable vacuum source (not shown) to retain a substrate 14 without exerting too much pressure thereon, since substrate 14 is often made of glass. A flood station 168 is provided which can expose the entire surface of substrate 10, when positioned beneath station 168 as illustrated in broken lines in FIG. 8, to a fluid typically used in the in situ process, and to which all features must be exposed during each cycle (for example, oxidizer, deprotection agent, and wash buffer). In the case of deposition of a previously obtained polynucleotide (such as a polynucleotide fabricated by the iterative sequence used in forming polynucleotides from nucleoside reagents on a support, as described above), flood station 168 may not be present.

A dispensing head 210 is retained by a head retainer 208. The positioning system includes a carriage 162 connected to a first transporter 160 controlled by processor 140 through line 166, and a second transporter 100 controlled by processor 140 through line 106. Transporter 160 and carriage 162 are used to execute one axis positioning of station 120 (and hence mounted substrate 10) facing the dispensing head 210, by moving it in the direction of arrow 163, while transporter 100 is used to provide adjustment of the position of head retainer 208 (and hence head 210) in a direction of axis 204. In this manner, head 210 can be scanned line by line, by scanning along a line over substrate 10 in the direction of axis 204 using transporter 100, while line by line movement of substrate 10 in a direction of axis 163 is provided by transporter 160. Transporter 160 can also move substrate holder 120 to position substrate 10 beneath flood station 168 (as illustrated by the substrate 10 shown in broken lines in FIG. 8). Head 210 may also optionally be moved in a vertical direction 202, by another suitable transporter (not shown). It will be appreciated that other scanning configurations could be used. It will also be appreciated that both transporters 160 and 100, or either one of them, with suitable construction, could be used to perform the foregoing scanning of head 210 with respect to substrate 10. Thus, when the present application recites "positioning" one element (such as head 210) in

relation to another element (such as one of the stations 120 or substrate 10) it will be understood that any required moving can be accomplished by moving either element or a combination of both of them. The head 210, the positioning system, and processor 140 together act as the deposition system of the apparatus. An  
5 encoder 130 communicates with processor 140 to provide data on the exact location of substrate station 120 (and hence substrate 10 if positioned correctly on substrate station 120), while encoder 134 provides data on the exact location of holder 208 (and hence head 210 if positioned correctly on holder 208). Any suitable encoder, such as an optical encoder, may be used which provides data  
10 on linear position.

Processor 140 also has access through a communication module 144 to a communication channel 180 to communicate with a distinct entity, e.g., a user or a system of the subject invention. Communication channel 180 may, for example, be a Wide Area Network ("WAN"), telephone network, satellite network,  
15 or any other suitable communication channel.

Head 210 may be of a type commonly used in an ink jet type of printer and may, for example, include five or more chambers (at least one for each of four nucleoside phosphoramidite monomers plus at least one for an activator solution) each communicating with a corresponding set of multiple drop dispensing orifices  
20 and multiple ejectors which are positioned in the chambers opposite respective orifices. Each ejector is in the form of an electrical resistor operating as a heating element under control of processor 140 (although piezoelectric elements could be used instead). Each orifice with its associated ejector and portion of the chamber, defines a corresponding pulse jet. It will be appreciated that head 210  
25 could, for example, have more or less pulse jets as desired (for example, at least ten or at least one hundred pulse jets). Application of a single electric pulse to an ejector will cause a droplet to be dispensed from a corresponding orifice. Certain elements of the head 210 can be adapted from parts of a commercially available thermal inkjet print head device available from Hewlett-Packard Co. as part no.  
30 HP51645A. Alternatively, multiple heads could be used instead of a single head 210, each being similar in construction to head 210 and being provided with respective transporters under control of processor 140 for independent movement. In this alternate configuration, each head may dispense a

corresponding biomonomer (for example, one of four nucleoside phosphoramidites) or an activator solution.

As is well known in the ink jet print art, the amount of fluid that is expelled in a single activation event of a pulse jet, can be controlled by changing one or more of a number of parameters, including the orifice diameter, the orifice length (thickness of the orifice member at the orifice), the size of the deposition chamber, and the size of the heating element, among others. The amount of fluid that is expelled during a single activation event is generally in the range about 0.1 to 1000 pL, usually about 0.5 to 500 pL and more usually about 1.0 to 250 pL. A typical velocity at which the fluid is expelled from the chamber is more than about 1 m/s, usually more than about 10 m/s, and may be as great as about 20 m/s or greater. As will be appreciated, if the orifice is in motion with respect to the receiving surface at the time an ejector is activated, the actual site of deposition of the material will not be the location that is at the moment of activation in a line-of-sight relation to the orifice, but will be a location that is predictable for the given distances and velocities.

The apparatus can deposit droplets to provide features which may have widths (that is, diameter, for a round spot) in the range from a minimum of about 10  $\mu\text{m}$  to a maximum of about 1.0 cm. In embodiments where very small spot sizes or feature sizes are desired, material can be deposited according to the invention in small spots whose width is in the range about 1.0  $\mu\text{m}$  to 1.0 mm, usually about 5.0  $\mu\text{m}$  to 500  $\mu\text{m}$ , and more usually about 10  $\mu\text{m}$  to 200  $\mu\text{m}$ .

The apparatus further includes a display 310, speaker 314, and operator input device 312. Operator input device 312 may, for example, be a keyboard, mouse, or the like. Processor 140 has access to a memory 141, and controls print head 210 (specifically, the activation of the ejectors therein), operation of the positioning system, operation of each jet in print head 210, and operation of display 310 and speaker 314. Memory 141 may be any suitable device in which processor 140 can store and retrieve data, such as magnetic, optical, or solid state storage devices (including magnetic or optical disks or tape or RAM, or any other suitable device, either fixed or portable). Processor 140 may include a general purpose digital microprocessor suitably programmed from a computer readable medium carrying necessary program code, to execute all of the steps

required by the fabrication station 38, or any hardware or software combination which will perform those or equivalent steps. The programming can be provided remotely to processor 141 through communication channel 180, or previously saved in a computer program product such as memory 141 or some other  
5 portable or fixed computer readable storage medium using any of those devices mentioned below in connection with memory 141. For example, a magnetic or optical disk 324a may carry the programming, and can be read by disk writer/reader 326. A cutter 152 is provided to cut substrate 10 into individual array units 15 each carrying a corresponding array 12.

10 The operation of the fabrication station will now be described. It will be assumed that a substrate 10 on which arrays 12 are to be fabricated, is in position on station 120 and that processor 140 is programmed with the necessary array layout information to fabricate target arrays 12 (sometimes referenced as the "target array layout" or similar). Using information such as the foregoing array  
15 layout and the number and location of drop deposition units in head 210, processor 140 can then determine a reagent drop deposition pattern.

Alternatively, the actual drop deposition pattern can be part of the array layout. In any event, the array layout can be provided to the fabrication station and communicated to memory 141 through communication channel 180. Processor  
20 140 controls fabrication, in accordance with the deposition pattern, to generate the one or more arrays on substrate 10 by depositing for each target feature during each cycle, a reagent drop set. Further, processor 140 sends substrate 10 to flood station 168 for intervening or final steps as required, all in accordance with the conventional in situ polynucleotide array fabrication process described  
25 above. The substrate 10 is then sent to a cutter 152 wherein portions of substrate 10 carrying an individual array 12 are separated from the remainder of substrate 10, to provide multiple array units 15. The foregoing sequence can be repeated at the fabrication station as desired for multiple substrates 10 in turn. In a variation of the foregoing, it is possible that each unit 15 may be contained with  
30 a suitable housing. Such a housing may include a closed chamber accessible through one or more ports normally closed by septa, which carries the substrate 10.

Following array fabrication, the fabricated array may then be forwarded, i.e., shipped, to the user using any convenient means. As such, following fabrication, one or more array units may then be forwarded to one or more remote customer stations.

5           Chemical arrays produced according to array layouts generated by the subject systems and methods find use in a variety of different applications, where such applications are generally analyte detection applications in which the presence of a particular analyte in a given sample is detected at least qualitatively, if not quantitatively. Protocols for carrying out such assays are well  
10       known to those of skill in the art and need not be described in great detail here. Generally, the sample suspected of comprising the analyte of interest is contacted with an array produced according to the subject methods under conditions sufficient for the analyte to bind to its respective binding pair member that is present on the array. Thus, if the analyte of interest is present in the  
15       sample, it binds to the array at the site of its complementary binding member and a complex is formed on the array surface. The presence of this binding complex on the array surface is then detected, e.g. through use of a signal production system, e.g. an isotopic or fluorescent label present on the analyte, etc. The presence of the analyte in the sample is then deduced from the detection of  
20       binding complexes on the substrate surface.

          Specific analyte detection applications of interest include hybridization assays in which the nucleic acid arrays of the subject invention are employed. In these assays, a sample of target nucleic acids is first prepared, where preparation may include labeling of the target nucleic acids with a label, e.g. a  
25       member of signal producing system. Following sample preparation, the sample is contacted with the array under hybridization conditions, whereby complexes are formed between target nucleic acids that are complementary to probe sequences attached to the array surface. The presence of hybridized complexes is then detected. Specific hybridization assays of interest which may be practiced using  
30       the subject arrays include: gene discovery assays, differential gene expression analysis assays; nucleic acid sequencing assays, and the like. Patents and patent applications describing methods of using arrays in various applications include: 5,143,854; 5,288,644; 5,324,633; 5,432,049; 5,470,710; 5,492,806;

5,503,980; 5,510,270; 5,525,464; 5,547,839; 5,580,732; 5,661,028; 5,800,992.  
Also of interest are U.S. Patent Nos.: 6,656,740; 6,613,893; 6,599,693;  
6,589,739; 6,587,579; 6,420,180; 6,387,636; 6,309,875; 6,232,072; 6,221,653;  
and 6,180,351. In certain embodiments, the subject methods include a step of  
5 transmitting data from at least one of the detecting and deriving steps, as  
described above, to a remote location.

Where the arrays are arrays of polypeptide binding agents, e.g., protein  
arrays, specific applications of interest include analyte detection/proteomics  
applications, including those described in U.S. Patent Nos.: 4,591,570;  
10 5,171,695; 5,436,170; 5,486,452; 5,532,128 and 6,197,599 as well as published  
PCT application Nos. WO 99/39210; WO 00/04832; WO 00/04389; WO  
00/04390; WO 00/54046; WO 00/63701; WO 01/14425 and WO 01/40803 – the  
disclosures of which are herein incorporated by reference.

As such, in using an array made by the method of the present invention,  
15 the array will typically be exposed to a sample (for example, a fluorescently  
labeled analyte, e.g., protein containing sample) and the array then read.  
Reading of the array may be accomplished by illuminating the array and reading  
the location and intensity of resulting fluorescence at each feature of the array to  
detect any binding complexes on the surface of the array. For example, a  
20 scanner may be used for this purpose which is similar to the AGILENT  
MICROARRAY SCANNER available from Agilent Technologies, Palo Alto, CA.  
Other suitable apparatus and methods are described in U.S. Patent Nos.  
5,091,652; 5,260,578; 5,296,700; 5,324,633; 5,585,639; 5,760,951; 5,763,870;  
6,084,991; 6,222,664; 6,284,465; 6,371,370 6,320,196 and 6,355,934. However,  
25 arrays may be read by any other method or apparatus than the foregoing, with  
other reading methods including other optical techniques (for example, detecting  
chemiluminescent or electroluminescent labels) or electrical techniques (where  
each feature is provided with an electrode to detect hybridization at that feature in  
a manner disclosed in US 6,221,583 and elsewhere). Results from the reading  
30 may be raw results (such as fluorescence intensity readings for each feature in  
one or more color channels) or may be processed results such as obtained by  
rejecting a reading for a feature which is below a predetermined threshold and/or  
forming conclusions based on the pattern read from the array (such as whether

or not a particular target sequence may have been present in the sample or an organism from which a sample was obtained exhibits a particular condition). The results of the reading (processed or not) may be forwarded (such as by communication) to a remote location if desired, and received there for further use  
5 (such as further processing).

The invention also provides programming, e.g., in the form of computer program products, for use in practicing the methods. Programming according to the present invention can be recorded on computer readable media, e.g., any medium that can be read and accessed directly by a computer. Such media  
10 include, but are not limited to: magnetic storage media, such as floppy discs, hard disc storage medium, and magnetic tape; optical storage media such as CD-ROM; electrical storage media such as RAM and ROM; and hybrids of these categories such as magnetic/optical storage media. One of skill in the art can readily appreciate how any of the presently known computer readable mediums  
15 can be used to create a manufacture that includes a recording of the present programming/algorithms for carrying out the above described methodology.

Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, it is readily apparent to those of ordinary skill in the art in light of the teachings of this  
20 invention that certain changes and modifications may be made thereto without departing from the spirit or scope of the appended claims.

WHAT IS CLAIMED IS:

1. A system for producing an array layout, said system comprising:
  - (a) a communication module comprising:
    - 5 (i) a input manager for receiving array request information from a user; and
    - (ii) an output manager for providing an array layout to a user; and
  - (b) a processing module, said processing module having an array layout developer configured to develop an array layout based on  
10 information that includes array request information received from a user, wherein said processing module is further configured to provide at least one of the following additional features:
    - (i) a collaboration manager configured to allow at least two  
15 different users to jointly provide array request information to said array layout developer;
    - (ii) a security manager configured to control information transfer in a predetermined manner between at least two different users via said system; and
    - 20 (iii) a vendor manager configured to provide access by a user to a service provided by at least one vendor.
2. The system according to claim 1, wherein said processing module is configured to include a collaboration manager configured to allow at least two  
25 different users to jointly provide array request information to said array layout developer.
3. The system according to claim 1, wherein said processing module is configured to include a security manager configured to control information  
30 transfer in a predetermined manner between at least two different users via said system.

4. The system according to claim 1, wherein said processing module is configured to include a vendor manager configured to provide access by a user to a service provided by at least one vendor.
- 5 5. The system according to claim 4, wherein said service comprises providing sequences of probes for one or more targets.
6. The system according to claim 1, wherein said communication module provides for remote communication between a user and said processor module.
- 10 7. The system according to claim 1, wherein said communication module provides for communication between a user and said processor module via the Internet.
- 15 8. The system according to claim 1, wherein said communication module provides for a graphical user interface (GUI) between a user and said processing module.
9. The system according to claim 1, wherein said system provides for at least
- 20 two of said additional features.
10. The system according to claim 9, wherein said system provides all of said additional features.
- 25 11. The system according to claim 10, wherein said array layout is a nucleic acid array layout.
12. A method of obtaining an array layout for a chemical array, said method comprising:
- 30 (a) inputting array request information into a system according to claim 1; and
- (b) obtaining an array layout from said system.

13. The method according to claim 12, wherein said inputting is via the Internet.

14. The method according to claim 12, wherein said inputting is via a graphical user interface.

15. The method according to claim 12, wherein said method further comprises ordering a chemical array having said array layout via said system from a chemical array vendor.

10

16. The method according to claim 15, wherein said method further comprises receiving said chemical array from said chemical array vendor.

17. A method of producing an array layout for a chemical array, said method comprising:

15

(a) receiving array request information into a system according to claim 1; and

(b) obtaining an array layout from said system to produce said array layout.

20

18. The method according to claim 17, wherein said receiving is via the Internet.

19. The method according to claim 17, wherein said receiving is via a graphical user interface.

25

20. The method according to claim 17, wherein said method further comprises producing a chemical array having said array layout.

21. The method according to claim 20, wherein said method further comprises shipping said chemical array to a customer.

30

22. A computer program product comprising a computer readable storage medium having a computer program stored thereon, wherein said computer program, when loaded onto a computer, operates said computer to:

- (a) receive chemical array request information from a user;
- 5 (b) develop a chemical array layout based on said received information; and
- (c) provide said array layout to said user;

wherein said computer program further controls said computer to perform at least one of the following additional tasks:

- 10 (i) allow at least two different users to jointly provide said chemical array request information;
- (ii) control system mediated information communication in a predetermined manner between at least two different users; and
- (iii) provide access by a user to a service provided by at least one
- 15 vendor.

FIG. 1

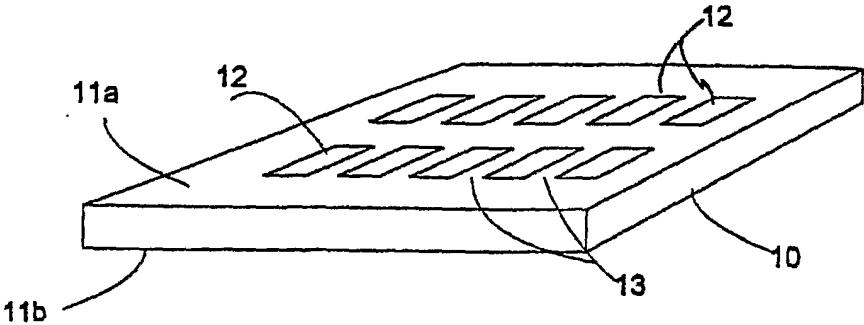


FIG. 2

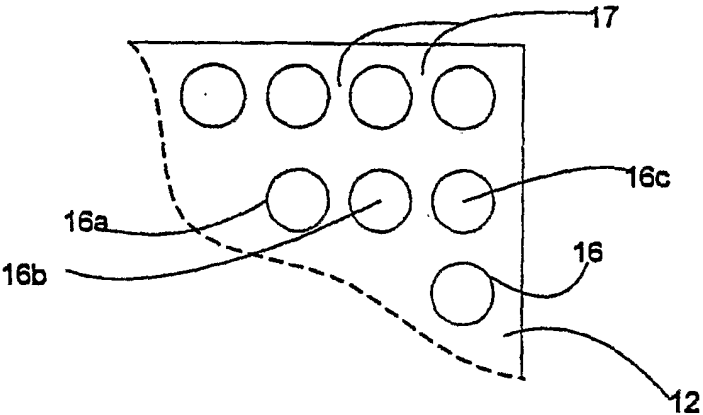
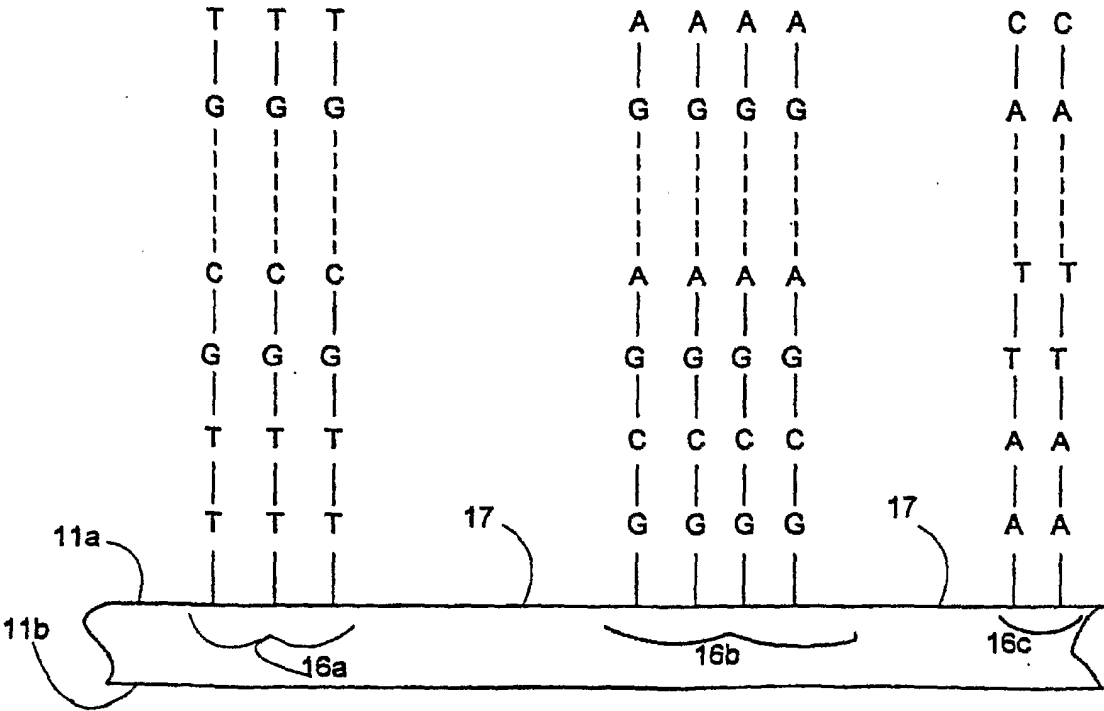
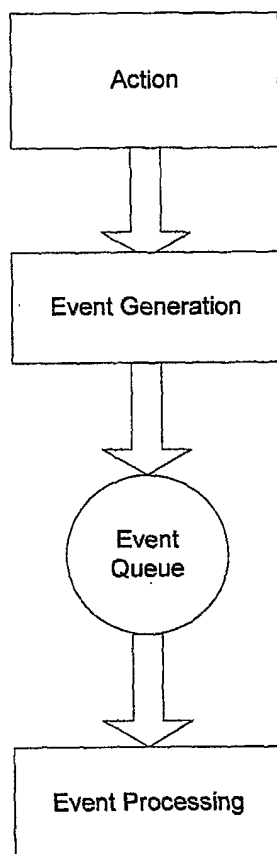


FIG. 3



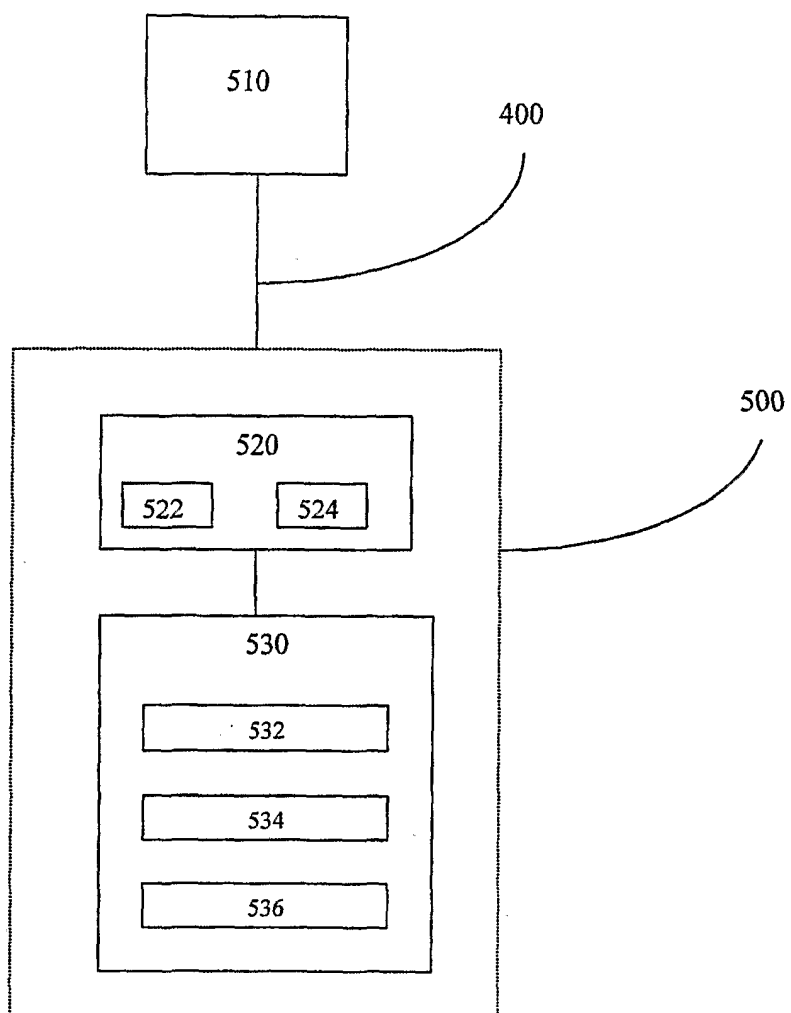
2/6

FIGURE 4



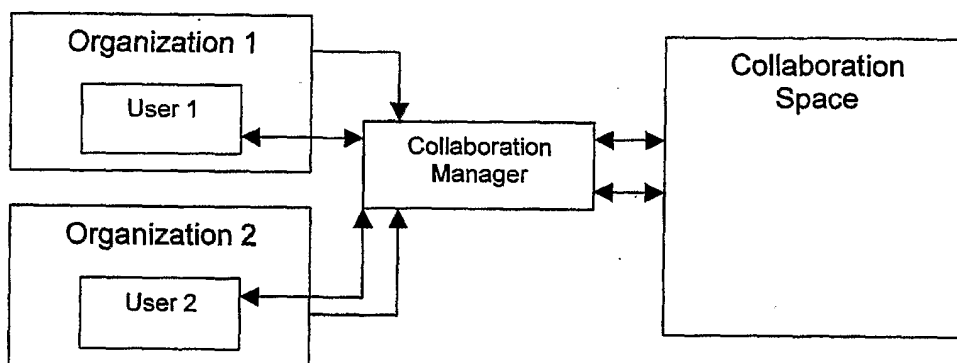
3/6

FIGURE 5



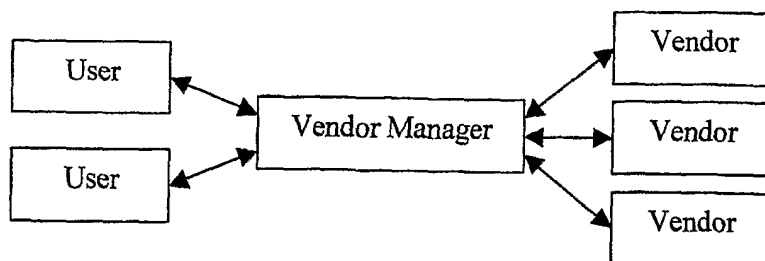
4/6

FIGURE 6



5/6

FIGURE 7



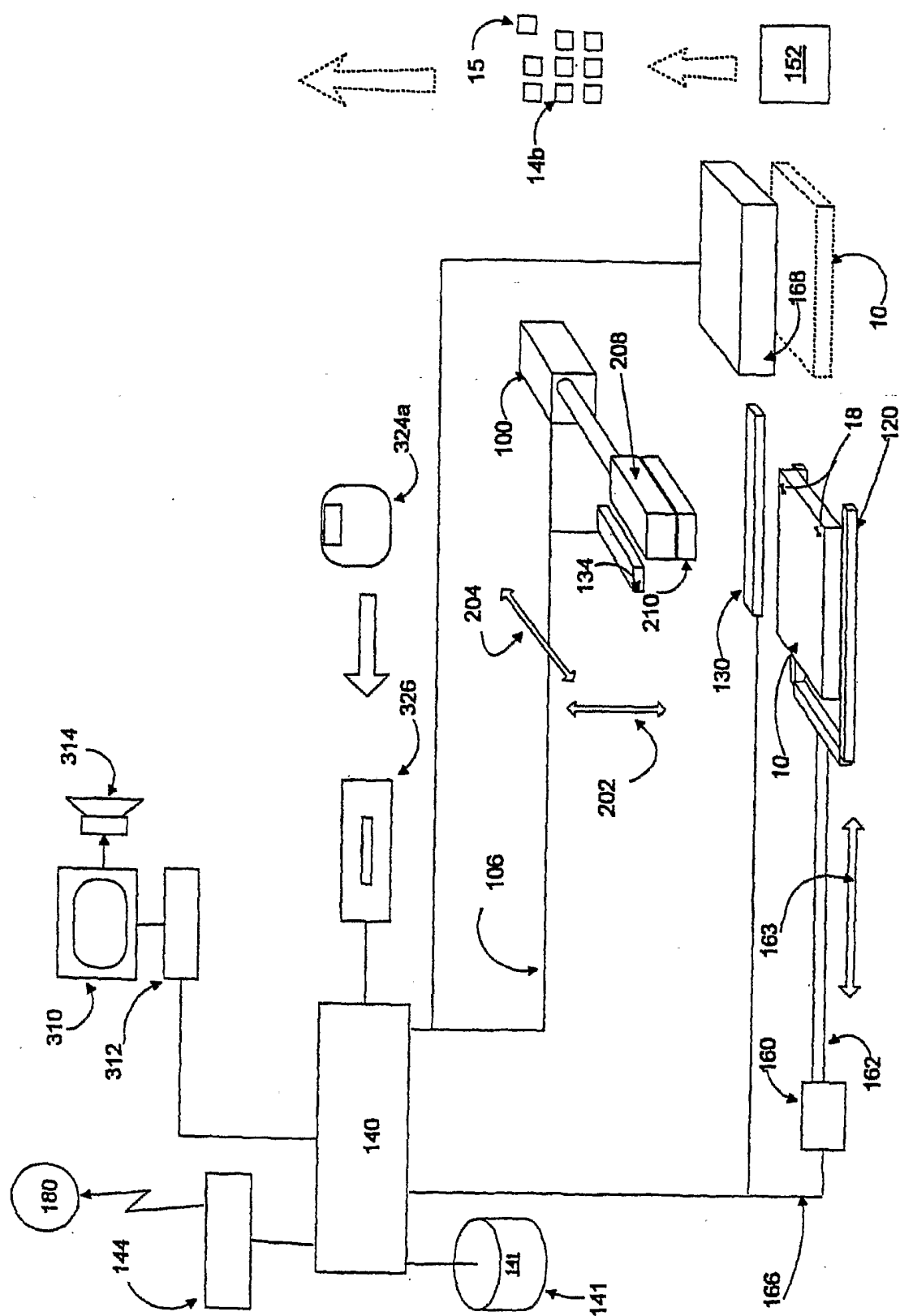


FIGURE 8

## INTERNATIONAL SEARCH REPORT

International application No

PCT/US2005/042003

A. CLASSIFICATION OF SUBJECT MATTER  
G06F19/00

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06F B01J

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data, PAJ, BIOSIS, EMBASE, INSPEC

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages                                                                                                                                                                                                                                             | Relevant to claim No. |
|-----------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------|
| X         | US 2003/120432 A1 (ZHOU XUE MEI ET AL)<br>26 June 2003 (2003-06-26)<br>the whole document<br>-----                                                                                                                                                                                                                             | 1-22                  |
| X         | US 2004/126840 A1 (CHENG JILL ET AL)<br>1 July 2004 (2004-07-01)<br>figures 4,5,7B,14,15<br>paragraph [0144]<br>-----                                                                                                                                                                                                          | 1-22                  |
| X         | WO 02/44994 A (THIRD WAVE TECHNOLOGIES,<br>INC; BROWER, AMY; BROW, MARY, ANN;<br>CRACAUER,) 6 June 2002 (2002-06-06)<br>page 100, line 25 - page 102, line 24<br>page 238, line 8 - page 242, line 17<br>page 244, lines 1-8<br>page 249, line 17 - page 256, line 18<br>page 290, line 15 - page 294, line 23<br>-----<br>-/- | 1-22                  |



Further documents are listed in the continuation of Box C.



See patent family annex.

## \* Special categories of cited documents:

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- "&" document member of the same patent family

Date of the actual completion of the international search

28 March 2006

Date of mailing of the international search report

04/04/2006

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

Godzina, P

## INTERNATIONAL SEARCH REPORT

International application No

PCT/US2005/042003

## C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages                                                                                                    | Relevant to claim No. |
|-----------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------|
| X<br>A    | WO 01/80155 A (COMBIMATRIX CORPORATION;<br>ANDERSON, BROOKE, P; QUARLES, PATRICK, A;<br>GHA) 25 October 2001 (2001-10-25)<br>page 26, line 20 - page 27, line 22<br>figure 1<br>----- | 1,4-8,<br>12-22<br>11 |
| X<br>A    | US 2002/075490 A1 (CHAPPELL JOHN C)<br>20 June 2002 (2002-06-20)<br>figures 1,2<br>paragraph [0041] - paragraph [0044]<br>-----                                                       | 1,4-8,<br>12-22<br>11 |
| X<br>A    | WO 01/69415 A (GENSET; PODDEVIN, BRUNO)<br>20 September 2001 (2001-09-20)<br>the whole document<br>-----                                                                              | 1,4-8,<br>12-22<br>11 |

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No  
PCT/US2005/042003

| Patent document<br>cited in search report |    | Publication<br>date | Patent family<br>member(s)                     | Publication<br>date                    |
|-------------------------------------------|----|---------------------|------------------------------------------------|----------------------------------------|
| US 2003120432                             | A1 | 26-06-2003          | NONE                                           |                                        |
| US 2004126840                             | A1 | 01-07-2004          | NONE                                           |                                        |
| WO 0244994                                | A  | 06-06-2002          | EP 1364334 A2                                  | 26-11-2003                             |
| WO 0180155                                | A  | 25-10-2001          | AU 6294601 A<br>CA 2377005 A1<br>EP 1312026 A2 | 30-10-2001<br>25-10-2001<br>21-05-2003 |
| US 2002075490                             | A1 | 20-06-2002          | NONE                                           |                                        |
| WO 0169415                                | A  | 20-09-2001          | AU 4099101 A                                   | 24-09-2001                             |