Accred Qual Assur (2007) 12:283–294
DOI 10.1007/s00769-006-0254-1

GENERAL PAPER

Defining and measuring biological activity: applying the principles of metrology

Craig M. Jackson · M. Peter Esnouf · Donald J. Winzor · David L. Duewer

Received: 11 November 2006 / Accepted: 28 December 2006 / Published online: 16 March 2007

Abstract A definition of biological activity is proposed that is superficially analogous to the equation relating the thermodynamic activity of a solute to its concentration via an activity coefficient. The biological activity of a molecular entity is defined as $A=cf$, where A is the activity, c the amount-of-substance concentration, and f is a parameter designated as ''inherent activity.'' Units and dimensions are determined by the type of activity, catalytic (katal) or binding (mol^{-1} L). The measurand is described by a chemical equation that identifies the entity for which an activity is being monitored. This definition of biological activity has the advantage of separating the chemical characterization of the entity in terms of structure and amount from the assessment of biological activity. Ideally, a homogeneous entity is used for the measurement of f. In instances where impure materials are used or the chemical equation defining the activity is unknown, the evaluated parameter should be designated as f' to denote its empirical nature. Any measurement of f or f' should be qualified with an appropriate estimate of measurement uncertainty.

Keywords Biological activity · Activity definition · Inherent activity

Introduction

Participation in biological processes is what makes particular molecular entities uniquely important and interesting[1]. However, the very importance of biological processes has fostered piecemeal approaches to the description of functional relationships between biological activities and the chemical substances that express them. This document attempts to define biological activity in a way that is both biologically informative and that enables development of the quantitative measurements needed to fully exploit the new knowledge. We identify the biological activity[2] of any entity by its ability to effect a change in a biological

C. M. Jackson (✉)
Hemosaga Diagnostics Corp., 5931 Seacrest View Rd.,
San Diego, CA 92121-4355, USA
e-mail: cjackso2@san.rr.com

M. P. Esnouf
Department of Clinical Biochemistry, Radcliffe Infirmary,
University of Oxford, Oxford OX2 6HE, UK

D. J. Winzor
Department of Biochemistry,
School of Molecular and Microbial Sciences,
University of Queensland, Brisbane,
Queensland 4072, Australia

D. L. Duewer
Analytical Chemistry Division,
Chemical Science and Technology Laboratory,
National Institute of Standards and Technology,
100 Bureau Drive, Gaithersburg, MD 20899-8390, USA

[1] The terms used here are intended to be consistent to the extent possible with current usage. New terminology is introduced when no existing term was considered appropriate. Terminology has been borrowed from other disciplines, and analogies with terms from those disciplines' lexicons have been employed to aid understanding the properties of the proposed definition for biological activity. Comments regarding the chosen terminology and selected definitions are provided in a Glossary at the end of this paper. References to these comments are indicated by Roman numerals in superscripts. It is recognized that some of the entries will be unnecessary for members of individual disciplines, but we are prompted to provide the Glossary by the unfamiliarity of the terminology across disciplines and the utility of having collected them in a single document.

[2] Biological activity, functional activity, and function are used as synonyms in this document.

process. We then devise a *framework* for defining and measuring biological activity based on the tenets of modern measurement science, but which, at the same time, is practical enough for use during the course of the discovery and characterization of new biologically active entities[3].

A fundamental dichotomy in the perspective regarding biological substances exists between chemistry and biology. In chemistry, an entity is identified by its molecular structure and the amount of the entity is typically measured in moles or grams. In biology, the ability of an entity to effect a change in a biological-process-based assay (bioassay)[i] identifies the entity. Quantification of the activity is obtained by an empirical dose–response relationship. More often than not, the dose is varied by diluting a sample of a biological material that includes the activity-expressing substance(s). While both disciplines may seek a biochemical-mechanism-based description of the process, a biologist is mainly concerned about what the entity does, whereas a chemist is mainly interested in what the entity is and how much of it is present. *Recognition of this dichotomy and the incorporation of substance, substance amount, and the property responsible for the expression of function into a new definition for biological activity is the primary goal of this document.*

The dichotomy between chemistry and biology has resulted in confusion regarding the measurements of biological entities and has confounded efforts to improve the comparability, traceability, and equivalence of the results of many biological assays. The quantitative measurement of biological materials by physical and chemical means is sometimes used to infer or predict biological activity, although the measurements themselves provide no information about the activity. For example, macromolecules such as proteins may be chemically measurable, but functionally inactive. Conversely, biological activity measurements are used to infer the amount of the entity that is present, but these measurements are fraught with potential for serious bias. Measurements made in the presence of interfering or inhibitory entities will frequently underestimate the amount of the biologically active entity. Similarly, an apoenzyme will be inactive in the absence of an obligatory cofactor. Activity measurements can also overestimate the amount of an associated molecular entity if other entities are present that enhance the activity, such as pro-enzyme activation. A distinction between biological and chemical entities has long been recognized in some World Health Organization (WHO) documents, where biological entities are specifically described as those that ''cannot be characterized adequately by physicochemical means alone'' [1]. While true, this stance offers no hint at a solution to the problem.

Strategy and goals for the proposed definition of biological activity

Our strategy for relating biological activity to the amount of a molecular entity begins by proposing a definition for biological activity that merges insights provided from both biological and chemical approaches to measurement. A common definition for biological activity is highly desirable for communicating information, particularly to those who make life-saving or life-threatening decisions on the basis of the reported values of markers of biological dysfunction.

To be most useful, the definition for biological activity should be applicable to both the simplest and the most complex reaction systems and molecules. It should also provide a means for refinement as knowledge of the biological process advances, a property we describe as extensibility[ii]. We implicitly use proteins as the archetype macromolecule; in subsequent documents, we will relate the definition to other macromolecules and to relatively low molecular mass entities. The value of this definition will be realized if, in medicine, it: (1) facilitates direct comparison of the potency (activity) of different biological entities that are being used therapeutically or being measured for their diagnostic value; (2) allows estimation of the extent of their equivalence; and (3) decreases the likelihood of incorrect diagnoses.

Several goals have been established for this definition of biological activity. The definition is intended first to provide a framework for communicating, discussing, and expanding knowledge of biological activity as it relates to chemical structure. A second goal is to separate the measurement of the structure-derived biological properties from that of the structure, physical properties, and amounts of the entities involved. In this regard, the definition follows the long-standing biochemical approach of relating structure and function. The third goal is to formulate a definition that is metrologically sound and, thereby, to facilitate use of measurements of biological activity without the confusion that is inevitable when arbitrary units or impermanent references are used. It is also intended that the definition be useful to bioinformatic efforts to make the rapidly expanding knowledgebase of biology more readily accessible.

[3] Proteomics-, metabolomics-, and bioinformatics-based inferences of function from structural homology, motifs, and domains in proteins are providing an unprecedented increase in the number of identified proteins, as well as descriptive information regarding relationships among proteins and the reactions in which they are involved. These new disciplines within biology, however, are, as yet, not very concerned with the elucidation, description, and measurement of the quantitative functional properties that are considered here.

The framework and a general definition for biological activity[iii]

Several truisms form the basis of the proposed definition. While perhaps ''obvious'' to some, we believe that these axioms establish the framework and the logic that underlies the definition. In a highly simplified way, the focus on a molecular entity with biological activity can be stated to be: (1) what it is; (2) what it does, where ''what'' is clearly plural; and (3) how much of it is present, both entity amount and expressed activity.

Axiom 1 Biological macromolecules are the predominant agents of biological activity. Many biological macromolecules express more than one definable activity or function.

Axiom 2 A particular function of a macromolecule is a property determined by structural attributes[iv] of the macromolecule.

Axiom 3 The function(s) of biological entities is (are) modifiable. These modifications may be the consequence of interaction(s)[v] with other molecules, the composition of the solution in which they are found, and temperature. Commonly recognized functional properties include: ligand-binding and binding site affinity, efficiency of expression of the activity, and specificity. An obvious example is the interaction of a protein with an allosteric effector.

Axiom 4 The biological activities of small molecules (ligands) are reciprocally related to the macromolecules to which they bind (acceptors). That is, the expression of a biological activity that can be ascribed to a small molecule is the consequence of the linkage[vi] between the small molecule bound to the macromolecule and the effect on the macromolecule [2, 3].

Axiom 5 Each distinguishable activity of a biological entity must be represented by the simplest possible set of chemical equations[vii]. Multiple chemical equations are expected for most macromolecules; complexities should be introduced parsimoniously.

Metrological principles

The goal of measurement science (metrology) is ''to achieve comparability of results over space and time'' [4]. The four interrelated measurement principles that form the basis for achieving comparability are: fitness for purpose, validation, uncertainty, and traceability. That is: a measurement system must be designed to provide measurement results of adequate quality for the task(s) at hand; the implementation of the design must be shown to indeed provide results of adequate quality; the measurement results must explicitly state what the expected measurement quality is; and the measurement results from a particular measurement system must be relatable to results obtained from other measurement systems through a common set of primary references.

Two tools of proven worth in the pursuit of comparability[viii] in physical and chemical metrology are the common system of units provided by the International System of Units[ix] (SI) and the common nomenclature provided by the International Vocabulary of Metrology (VIM) [5]. These internationally accepted systems of units and vocabulary facilitate the achievement of measurement comparability and the evaluation and description of the extent of its achievement.

The rigorous application of these principles and tools to the measurement of biological activity has been limited to small molecules and ''procedure-defined measurands''[x], e.g., enzymes [6]. We believe that these principles can be applied generally to the measurement of other biological entities, including macromolecules, and can be directly related to their activity in biological systems. The proposed definition for biological activity provides a means for applying metrology to all biological substances.

The algebraic definition of biological activity

The first step in applying metrological principles to the measurement of a biological entity requires the separation of entity and entity amount from the entity's expression of function. Although it has always been evident that biological activity is dependent on what the entity is and its amount, explicit separation of amount and activity is not commonly made. Because such a separation cannot be made when activity is assigned in arbitrary units to a complex mixture, one component of which is assumed to express the activity, this ''traditional'' approach will always be severely limited. We propose a parameter f which links entity, entity amount, and biological activity to achieve this separation.

Based on the axioms stated above and an imperfect analogy with thermodynamic activity[4], the following simple algebraic equation is proposed to define the biological activity of an entity:

[4] Individuals with backgrounds in physical chemistry will immediately recognize this as an analog of the equation that relates the thermodynamic activity of electrolytes via the product of concentration and an activity coefficient. Use of this analogy, as extensively discussed by Polya [7], is a heuristic tool that can facilitate discovery, analysis, and refinement of a concept through the process of evaluating applicability (strengths) and limitations (weaknesses).

$$A = cf$$

where A is the biological activity, c is the amount-of-substance concentration of the entity of interest, and f is a parameter designated as *inherent activity*[xi]. The description of f as ''inherent'' can be legitimately applied without ambiguity only for an idealized reference material in which all of the molecular entities present are identical[xii]. However, materials that approximate this requirement are rapidly becoming available[5].

This equation emphasizes that the measurement of the concentration alone does not suffice to describe the functional capability of a biological entity, nor does the measurement of activity alone infer unambiguously the entity concentration.

With regard to this definition, the aforementioned biological perspective may be regarded as a focus on A without recognizing it as the product of c and f, whereas the chemical perspective is focused on c without considering that the functional significance of c lies in the value f. The proposed definition enables harmonizing the two perspectives by separating the chemical variable and biological variable into two terms that, only together, disclose functional activity.

The definition of biological activity, $A=cf$ describes a linear dose–response curve of biological activity as a function of concentration. The intercept on the ordinate is zero (0). Ideally, concentration is an entity amount traceable to a certified reference material (CRM) of high purity.

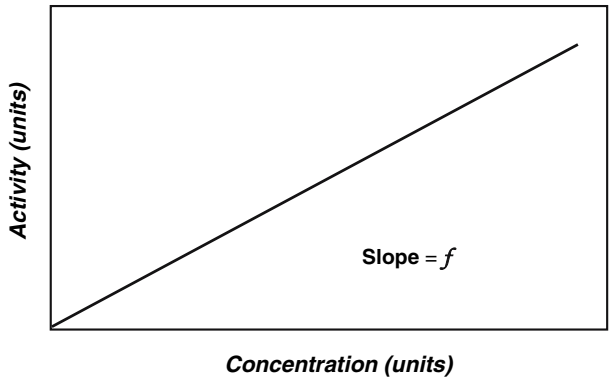

Fig. 1 Dependence of biological activity on concentration

When no suitable CRM is available and a suitably stable and homogeneous reference material cannot be identified or developed, f' is used to indicate that the relationship is empirical (Fig. 1).

When the entity concentration c is known, the definition describes a straight line that passes through the origin and has a slope f. The parameter f is a property of the biological entity for which the defined function is being measured and contains the relevant information about the ability of the particular attribute of the molecular entity to express its activity. Although the equation itself does not demand that the concentration be expressed as mol L^{-1}, inferences about the function that can be related to molecular structure follow from concentrations expressed on that scale. In this regard, information related to the structure of the molecular entity and its related properties (e.g., molecular mass) is intended to be linked to c without reference to the function(s) expressed by the entity.

It must be noted that attributes (e.g., sites) of biological macromolecules limit the site occupancy to the concentration of the ligand that binds to the site, its stoichiometry, and the binding constant of the site for the ligand. Consequently, the process of binding is intrinsically non-linear and is commonly described by a rectangular hyperbolic function. A strictly linear dependence of A upon c is, thus, only observed in the limiting slope as c approaches zero. In this regard, f differs from the classical activity coefficient, which approaches one as c approaches zero.

An additional, and possibly the greatest, advantage of this definition is that the variable f can be interpreted using contemporary knowledge of biochemical reaction mechanisms[6]. In fact, it is the ability to interpret f mechanistically that makes it useful for harmonizing the biological and

[5] An idealized macromolecule might be considered to be analogous to an ideal gas, but the focus is on the structural uniformity of the chemical species, rather than the absence of intermolecular interaction. However, it is useful to envision the idealized macromolecule as analogous to a molecule in a vacuum—without external entities or forces to act on it. In this sense, the idealized macromolecule is considered to be free of a matrix and the influence of entities that comprise the matrix. Although not realizable in practice, this formulation simplifies discussion because it separates matrix component influences from structure-based variations that change the inherent activity of the macromolecule. A protein is considered to be a single chemical species based on its being the product of transcription of a definable exon. All polymorphs must be considered to be separate chemical entities. A polymorphism that changes an attribute in a discernible way can be related to the corresponding change in the inherent activity. No detectable change may be a common occurrence; in such situations, the polymorphism is linked to c in the defining equation and the value of f is noted as not being demonstrably different. A polymorphism can be without effect on one defined biological activity but with significant effect on another; an example of pleiotropy, multiple expressions of an alteration in a single attribute. The most likely practical realization of the idealized protein is a recombinant product with experimentally demonstrable homogeneity of amino acid sequence and post-ribosomal modification. A clear advantage of such a realization of the entity is that, when such a protein reference material can be produced in sufficient quantity, it can be used as the calibrator for the measurement of both entity amount and biological activity.

[6] Analogy with the thermodynamic activity coefficient and the theoretical models that have been developed to interpret the activity coefficient is the underlying basis for extending the interpretation of f via biochemical models, e.g., enzyme kinetic and ligand acceptor models.

chemical perspectives. It is, moreover, the capability of the mechanistic descriptions of the biological process, e.g., kinetic equations describing enzymatic reactions, that enables inherently non-linear processes to be transformed into forms that permit such simple description.

For f to be interpreted as an inherent activity of the macromolecule, the process that is being measured must be the process described by the defining chemical equation. When a measurement procedure is or cannot be limited exclusively to a single attribute, separate chemical equations must be written to describe each attribute and reaction that occurs in the measurement procedure in order to avoid confusion in interpreting f. If the measured biological activity comprises the expression of multiple functions, then f is unlikely to be related to a single or particular attribute of the macromolecule in a readily discernible way, unless all of the individual functions are specified and their combined effects are taken into account. Succinctly, the utility of the proposed definition demands appropriate definition of the measurand.

In situations where the definition seemingly suffices, but doubt remains about the adequacy of the chemical equation(s) used to define the function, the inherent activity should be defined as f' to indicate the doubt. This is expected to be the prevalent situation. If, as is inevitable in the earliest stages of discovery and characterization, the measurement procedure is necessarily empirical, then the functional capability parameter should again be designated as f' to signify its lack of a proven chemical basis. As information regarding the entity and its measurement increases, refinements to the definition of the function, the measured value, and the extent of interpretation of f can be made. Moreover, when the limitations of the initial estimate for f are stated, the changes and the causes that demand change can be recognized and insight gained from the refinement process. The approach based on this proposed definition promotes meeting the goals of metrology by acknowledging uncertainty, first qualitatively and subsequently quantitatively, and in any particular measurement procedure, it fosters a clear definition of the measurand, the effects of interactants and influence quantities, and a rational discussion of fitness for purpose.

Although it is universally recognized that biological activity does not appear in a hypothetical isolated state, i.e., as a description only on paper or as a depiction of the 3D structure of a macromolecule, this seems to be easily forgotten when actually describing a measurand. Biological activity is a reflection of interactions between molecules noted in the defining equation(s) and their transformation(s) in the milieu provided by the medium (solution) in which the reactions occur. Further, in this regard, interaction between a substrate and an enzyme or a ligand and a receptor implies a reciprocal relationship that requires the consideration of both entities. Each molecular entity will possess its own attributes and their associated inherent functional capabilities that are represented by the value of f; but for simple low molecular mass entities, f may simply be unity. The linkage relationship is analogous to the linked functions of thermodynamics [2, 3] and reciprocal relationships of thermodynamics.

Entity identity, entity amount, structural attributes, and the expression of activity—practical considerations for the measurement and interpretation of f

The separation of biological activity into chemical and biological terms via c and f forms the basis for resolving the dichotomy of chemical and biological perspectives. In this regard, the differences in the chemical and biological perspective to which we attribute much of the confusion surrounding the measurement of biological activity now become the means for eliminating that confusion.

Macromolecular entities such as proteins can be treated as chemical entities in the same ways as simpler molecules. Although macromolecules are much more complex, technological advances now make it feasible to approach their characterization in the same ways as is done for relatively low molecular mass entities. The characterization of protein molecules and the determination of their concentrations in the SI unit of mol L^{-1} is becoming practical; many commercial, governmental, and academic organizations exist to provide the necessary measurements of amino acid composition, post-ribosomal modification, prosthetic group content, etc. The measurement of c for proteins can be rigorous, although the uncertainty of the measured value can be very large because of calibration bias, unless the procedure used is directly related to the particular protein and its amino acid composition. However, when accompanied by an appropriate uncertainty estimate[xiii], c will be suitable for use in this definition for biological activity, in part, because the approach anticipates refinement.

Estimation of the uncertainty of c must recognize and consider the intrinsic heterogeneity of biological macromolecules. Macromolecules that have been derived from biological sources, even highly purified preparations without detectable contaminants or observable heterogeneity, will inevitably be heterogeneous to some extent. Materials of biological origin are heterogeneous because of genetic mutations, polymorphisms, and variability in post-translational modification. In practice, this is a consequence of the pooling of tissues (e.g., blood plasma) from multiple individuals prior to purification of the entity. Such micro-heterogeneity can be very small, and, in many cases, of little or no importance. However, values for the mea-

sured properties of these ''real'' materials are averages of the properties of the individual molecules.

Failure to recognize such sources of heterogeneity can lead to unproductive discussion and inappropriate estimation of the true uncertainty of the inherent activity f to be assigned to the reference material. For this discussion, however, it is conceptually helpful to assume a hypothetical reference material that consists of molecules, all of which are identical in structure. Such an idealized macromolecular entity provides an advantage in that the functions of the macromolecule can be directly related to its sequence and sites on the macromolecule that are altered post transcriptionally or translationally. Moreover, without unnecessary discussion, the structures can be assumed to be represented by their primary sequences and 3D structures. A real example that may approximate the idealized reference material is a recombinant protein produced in a system with complete fidelity in transcription, translation, and post-translational modification. In the context of metrological traceability, such an idealized material might be considered to be an approximation to a primary, pure-entity reference material. Entity identity is thus defined, and the entity amount and the uncertainty of its measurement is included in the assessment of c in the defining equation. All secondary batches of the entity are traced to the primary material through the values of the substance amount, the variations in the nature, and the extent of post-translational modifications and the effects of suchlike on the value of f for the batch.

It is conceptually important to distinguish intrinsic heterogeneity as described above from heterogeneity resulting from the reference material of biological origin being a mixture of different entities. The consequences of this latter type of heterogeneity, i.e., the presence of other entities (contaminants) in mixtures of biological materials, such as blood plasma, must be treated with regard to the influences they exert on the measurements of the biological activity of the idealized reference material or its practical equivalent. Because some such entities, isolated in conjunction with the macromolecule, are likely to be interactants (activity modifiers), they must be explicitly measured whenever they are known. Matrix-based reference materials in which the measurand is a pure entity can be characterized with respect to the effects of matrix components on the measurand, whereas reference materials in which activity is defined by the measurand native to the matrix cannot be so characterized and, thus, are predisposed to the effects of interactants and influence quantities that are probably unknowable.

The specific interpretation of f is intentionally and obligatorily linked to the process and chemical equations that define and describe the function. As previously noted, the definition for A is analogous to the definition of activity (a) in classical thermodynamics, but with appropriate caveats. Thus, f can be considered to be analogous to γ, the thermodynamic activity coefficient. Interpretation of γ for ions is based on the Debye–Hückel–Onsager theory of electrolytes [8, 9][7]. In the context of enzyme biological activity, the obvious model for interpreting f is the Michaelis–Menten equation (including its forms that include the effects of modifiers).

By way of illustration that this is an established approach for characterizing enzymes and their activity, the chemical equation that describes the catalytic action of an enzyme in its simplest form is:

$$\text{Enzyme} + \text{Substrate} \overset{K_\text{m}}{\leftrightarrows} \text{Enzyme:Substrate}$$

$$\text{Enzyme:Substrate} \xrightarrow{k_\text{c}} \text{Enzyme} + \text{Product}$$

The two processes involved in the expression of enzyme activity are substrate recognition and binding, represented by K_m (the Michaelis constant), and chemical transformation, represented by k_c. Interpretation, thus, links biological activity (A) to the concentration of the enzyme, c_enzyme, from:

$$A = -\frac{\mathrm{d}c_\text{substrate}}{\mathrm{d}t} = c_\text{enzyme} k_\text{c} \frac{c_\text{substrate}}{K_\text{m} + c_\text{substrate}}$$

A linear dependence on c_enzyme is usually observed, but there are situations in which this generalization does not hold[8]. If the independent variable is $c_\text{substrate}$, then linearity is only apparent when $c_\text{substrate} \ll K_\text{m}$. The selection of appropriate conditions for using an enzyme-catalyzed reaction as a means for the measurement of c_enzyme, $c_{\text{substrate}>K\text{m}}$, and for $c_\text{substrate}$, $c_\text{substrate} \ll K_\text{m}$, are well and long established [10]. The single parameter that informs the catalytic efficiency of an enzyme for the particular substrate is the ratio of k_c/K_m; conventionally, this is given with dimensions $\text{mol}^{-1}\ \text{L}\ \text{s}^{-1}$ or a decimal multiple, e.g., $\mu\text{mol}^{-1}\ \text{L}\ \text{s}^{-1}$. Since 1999, the SI name for the unit of catalytic activity (A) is the katal (kat), having dimensions s^{-1} mol. The SI expression for catalytic activity concentration is katal per cubic meter, kat m^{-3}. Based on

[7] Although formally based on the simple equation of thermodynamics and the interpretation of the activity coefficient for electrolytes in the Debye–Hückel–Onsager theory, the origin of corrections for non-ideality, solution composition, and solvent interactions dates to the van der Waals corrections to the ideal gas law and the application of statistical mechanics to complex systems by McMillan and Mayer and many others.

[8] Any of the many texts and monographs on enzyme kinetics cover the variety of enzyme mechanisms and the kinetic equations that provide interpretations of f. It should be noted that it is recognized that K_m and k_c are not so simply interpreted in most situations as implied here.

the SI unit, the katal, the units for f are L s^{-1}. The efficiencies of enzyme-catalyzed processes and the low concentrations of the enzymes and substrates in such reactions lead to measured values reported in μkat L^{-1} or nkat L^{-1}.

A similar chemical equation would apply for a system consisting of a macromolecular receptor and a small drug molecule that alters the biological behavior of the receptor. Likewise, if the molecular entity were an oligonucleotide, then binding alone might be the biological activity. However, if oligonucleotide binding were the first step in the process of gene transcription, then additional chemical equations and reactions would be included.

Discussion

The results from the measurement of entities of biological origin constitute a substantial part of the information upon which medical diagnoses are made. Moreover, it has been estimated that 60–70% of medical diagnoses are based on laboratory tests [11]. Assays, or more specifically, measurement procedures, provide values of biomarkers that underpin diagnostic decisions and therapeutic interventions. The comparability and equivalence of measurement results for biomarkers are necessary to support appropriate diagnoses and avoid misdiagnoses.

The utility of genomic information ultimately derives from its ability to predict phenotype as exhibited in resistance or susceptibility to disease. Although historically applied to a trait exhibited by an organism, the concept of phenotype can be, and now is, applied to both cells and macromolecules. Phenotype expression by a macromolecule is, in essence, what we call biological activity. Genetic mutations and polymorphisms can, and frequently do, alter function, thus, the need for measurement of the functional consequences of the structural changes.

The complexity of biological systems and biological macromolecules creates a formidable challenge for the development of traceable reference materials and procedures that provide equivalent values for the measurand. This is particularly so for proteins because of the multiple functions that a single protein frequently expresses, and also because of the presence in the biological sample of different proteins that nominally express the same or a similar function. A practical but limited solution to this problem has been achieved for the measurement of enzymatic activity under highly optimized and rigidly defined conditions, a solution that designates enzymatic activity measurements as catalytic activity concentrations and procedure-defined measurands [6][x]. The ''procedure-defined measurand'' approach does not lend itself to the acquisition of new information about the function of the measurand and does not address the challenge that originates from the need for translating genotype to phenotype at the level of the molecule and its structure.

The proposed definition can also be viewed as an attempt to improve on the ''procedure-defined measurand'' approach; we have devised a simple equation that separates the more definable entity and entity amount, the pertinent parameters as perceived by a chemist or chemical metrologist, from the expression of the biological function of the entity. To do so, we have further attempted to take advantage of much that is already known. However, at this stage, the approach has not been applied generally to the measurement of entities of biological origin and their properties that imbue them with biological function[9].

A benefit that derives from the proposed description of biological function and the particular structural attribute responsible extends beyond the definition of the measurand. The proposed definition can also focus attention on the fitness for purpose of the measurement procedure[10]. Functions, reactants, and interferents that cannot be described completely may well be important in identifying the limitations of particular measurement procedures. Although incompletely described procedures might be suitable for some purposes, perhaps because there are no alternative procedures, they will be unfit for reference measurement procedures that are intended to be included in a metrological traceability chain.

The utility of a general definition extends beyond the measurement of biomarkers used in medical diagnosis. The estimation of the potency and, thus, dose of therapeutic agents derived from biological materials is another obvious area of applicability. If interactants and interferences are identified during the development and validation of a measurement procedure based on the general definition, the procedure should prove beneficial for the monitoring of therapeutic, as well as adverse, effects of drugs.

As already noted, the lack of comparability and absence of a high degree of equivalence of biological activity measurements create opportunities for inappropriate diagnostic inferences to be made. When different measurement procedures (e.g., different manufacturers' kits for nominally the same entity or activity) do not produce equivalent results, the clinician can be misled. Similarly, when the results

[9] It may be argued that classical enzymology has, for decades, employed this approach to relating biological (catalytic activity) to structure and amount of substance; regrettably, enzymology and enzyme kinetic approaches have suffered neglect because of the excitement of genomics and related areas of science.

[10] The capability of this definition is, in one sense, an implementation of Popper's [12] falsifiability criterion for testing scientific hypotheses. The chemical equations describe the hypothesis (model) upon which the measurement procedure is based. The mechanistic equations provide the predicted behavior and the realization of the definition in the correspondence between the predicted and observed behavior.

of large, multiple-site epidemiological studies are interpreted, bias can obscure relationships and also lead to invalid inferences. The proposed definition for biological activity attempts to reduce the incidence of judgment errors that are derived from the ambiguity and inherent complexity of the measurement of biological activity and biological entities. It does so by separating the strictly chemical properties of a measurand from the functional consequences of those properties—biological or, more strictly, biochemical function. Moreover, because the definition is intended to be simple yet rigorous, its application can be used to identify and correct for interferents and influence quantities that could otherwise go unnoticed because this definition embodies predictability and testability [12].

The authors recognize that all change is, itself, initially prone to be a source of confusion and mistakes. In an attempt to minimize this and also to avoid needless change, we have drawn on the insights of individuals whose interests and contributions to laboratory medicine derive from biological perspectives, as well as from those whose insights are based on the measurement of simpler chemical entities. Similarly, we have adapted, by analogy, a definition for biological activity that has a precedent in chemical thermodynamics—the concept of chemical concentration and an activity coefficient that can be interpreted in molecular terms to explain why the behavior of even simple molecular entities is not always compatible with 100% efficiency of function. To the extent that readers of this proposed definition recognize familiar principles being applied to a more complex system, we hope that this approach will have a good chance of achieving its intended effect.

On a more technical level, the most valuable aspect of the proposed definition derives from its separation of the ''chemical'' and ''biological'' aspects of the defining chemical equation(s). This in itself is not novel, but neither the utility nor even the necessity of this separation seems to have been recognized. Inferences from the measurement of activity regarding entity amount or activity from the measurement of entity are now clearly related to a defined parameter, *f*. Because the value for *f* is related to the composition of the solution in which the measurement procedure occurs, it is constrained to be measured under well-controlled conditions—a long-recognized prerequisite for all procedure-defined measurands. Moreover, as illustrated by the Michaelis–Menten equation for an enzyme, *f* is interpretable in the same fashion as charge effects on ion properties in the Debye–Hückel–Onsager theory for electrolytes. The appropriate mechanistic equation is determined by the chemical equations that describe the biological activity.

The definition also aids in harmonizing the empirical and mechanistic interpretation of *f* because of its simple linearity, a feature that, for intrinsically non-linear biological processes and phenomena, can be obtained by mathematical transformation of the mechanistic equation(s) or, as already noted, by extrapolation of *c* to zero. However, the behavior of the function that mechanistically defines biological activity does not have to be linear, but it must be explicitly stated. Another evident benefit of this simple definition is that it is consistent with practices in pharmacology and other biological fields, although some descriptive terms used in other areas may not be strictly equivalent to *f*.

The focus on attributes as ''agents'' that produce or express function and the requirement that the attribute be included in the chemical equation should be useful in databases that catalog properties of biological macromolecules. The separation of *c* and *f* provides for extensibility; structure and substance amount are included in *c* and the expressed function in *f*. Structural attributes are related to motifs, domains, and other descriptors of structural features of macromolecules, particularly proteins. The prescription of an ''idealized macromolecule,'' moreover, provides a path between the entity and its 3D structure as provided by crystallographic and magnetic resonance methods. Such a path is, in our minds, necessary if the information from all disciplines is to be used to devise and develop measurement procedures that are suitable for all intended purposes. When limitations exist, they can be known and be used as caveats to prevent inappropriate inferences and decisions.

Acknowledgments The authors thank Paul De Bièvre for his suggestions and helpful comments, Helen Parkes of LGC, Ltd., for calling our attention to the WHO statement regarding biological substances [1] and the Department of Pharmacology and Experimental Therapeutics, Boston University, for permission to quote from the department's website.

Glossary

Definitions of terms from official sources are italicized to distinguish them from those terms as used in this document that may deviate from the official definitions. This glossary also provides a commentary on the selection of terms and the perspective that underlies the proposed definition of biological activity.

(i) **Bioassay**, **assay**, **test**, and **measurement procedure** are terms which are sometimes synonymous, but are used in different contexts, by different disciplines, and with different levels of rigor in their descriptions. It is common in complex bioassays to relate the response to the amount of dosing substance using a linear function that may not be related to any particular molecular mechanism.

– *Bioassay* [A1]: a procedure for determining the concentration or biological activity of a substance (e.g., vitamin, hormone, plant growth factor, antibiotic) by

measuring its effect on an organism or tissue compared to a standard preparation.

- *Bioassay or biological assay* [A2]: ''The determination of the potency of a physical, chemical or biological agent, by means of a biological indicator... The biological indicators in bioassay are the reactions of living organisms or tissues.'' Principles characterizing the potency of a bioassay include:
  - Potency is a property of the material to be measured, e.g., the drug, not a property of the response. Ordinarily, the relationship between changes in behavior of the indicator and differences in drug dose (a dose–effect curve) must be determined as a part of each assay.
  - Potency is relative, not absolute. The potency of one preparation (the ''unknown'') can be measured only in relation to the potency of a second preparation (the ''standard'' or ''reference drug'') that elicits a similar biologic response. When the absolute amounts of standard used in the assay are known, the results of the assay can be used to estimate the amount—in absolute units—of biologically active material contained in the unknown preparation.
  - A bioassay provides only an estimate of the potency of the unknown; the precision of the estimate should always be determined by using the data of the assay [A3].

(ii) **Extensibility**. ''In software engineering, extensibility (sometimes confused with forward compatibility) is a system design principle where the implementation takes into consideration future growth. It is a systemic measure of the ability to extend a system and the level of effort required to implement the extension. Extensions can be through the addition of new functionality or through modification of existing functionality. The central theme is to provide for change while minimizing impact to existing system functions'' [A4].

(iii) **Biological activity** as defined in this document is intentionally related to the use of the term in other contexts and standards to the greatest extent possible. The definition proposed here focuses on its application to macromolecules of biological origin (''biologicals''), but can be applied to smaller molecules in the same manner.

- *Biological activity* [A5]: the biological activity is the specific ability or capacity of a particular molecular entity to achieve a defined biological effect. Potency is the quantitative measure of the biological activity.
- *Potency* [A5]: as applied to the biologics and biotech areas, this is the measure of the biological activity using a suitable quantitative biological assay (also called potency assay or bioassay), based upon the product's attribute, which is linked to the relevant biological properties.
- *Potency* [A2]: an expression of the activity of a drug, in terms of the concentration or amount needed to produce a defined effect; an imprecise term that should always be further defined.
- *Potency* [A6]: ''Having great power, influence, or effect.''
- *Intrinsic efficacy*, *intrinsic activity* [A2]: the property of a drug that determines the amount of biological effect produced per unit of drug-receptor complex formed. Two agents with different intrinsic activity combining with equivalent sets of receptors may not produce equal degrees of effect, even if both agents are given in maximally effective doses. In such a circumstance, the agent producing the greater maximum effect has the greater intrinsic activity. Intrinsic activity is not the same as ''potency,'' and may be completely independent of it.

(iv) **Attribute**: a quality or feature regarded as characteristic or inherent [A5]. Examples of attributes of macromolecules are: ligand binding sites, catalytic sites and site residues, sequences of residues, post-ribosomal modification sites, epitopes, or other structurally definable entities. A change in the structure of a protein as the result of a mutation or site-directed mutagenesis is considered to constitute an attribute.

(v) **Interactant(s)** [A1]: substances that bind and alter the properties of the macromolecule for which measurements of function are being performed. Interactants may include any substance, including ionic substances. However, designation as an interactant is considered to be more specific than electrostatic screening by ions. The term has been used also in the biochemical literature in a more restricted sense, e.g., the description of energy transfer between donors and acceptors that is measurable by fluorescence.

(vi) **Linkage** is a thermodynamic concept that indicates the relationship between an interactant and a macromolecule to which it is binding, e.g., the interaction between ligand and receptor or protonation of hemoglobin and the Bohr effect [2, 3].

(vii) **Chemical reaction equation and chemical reaction** [A1] are terms used to achieve clarity and to be consistent with the definitions of the IUPAC as quoted below.

- *Chemical reaction equation* [A1]: symbolic representation of a chemical reaction, where the reactant entities are given on the left hand side and the product entities on the right hand side. The coefficients next to the symbols and the formulas of entities are the absolute values of the stoichiometric numbers. Different symbols are used to connect the reactants and products with the following meanings: =for a stoichiometric relation; $\rightarrow$ for a net

forward reaction; ⇄ for a reaction in both directions; H for equilibrium.

– *Chemical reaction* [A1]: a process that results in the interconversion of *chemical species*. Chemical reactions may be *elementary reactions* or *stepwise reactions*. This definition includes experimentally observable interconversions of conformers. Chemical equations can also describe an interaction without chemical change, i.e., binding that may or may not detectably alter the conformation or transport functions of the macromolecule.

(viii) **Comparability** is defined differently in the metrology literature and official documents of other international regulatory or advisory organizations and in common usage. Comparability is frequently used according to a generic definition, ''able to be likened to another, similar, of equivalent quality'' [A5], but with similarity in magnitude implied by the context in which the term is used. The metrologically restricted definition avoids this ambiguity and uses extent (degree) of equivalence to indicate similarity in a specified property. In this document, the parameter *f* reflects the extent of equivalence of the functional capability and is intended to be consistent with the metrological terminology.

– *Comparability of measurement results* [5, A8]: ''...quantities of the same kind that are metrologically traceable to the same reference.''
– *Comparability* [11, A9]: The demonstration of comparability does not necessarily mean that the quality attributes of the pre-change and post-change products are identical, but that they are highly similar and that the existing knowledge is sufficiently predictive to ensure that any differences in quality attributes have no adverse impact upon safety or efficacy of the product.

The term commutability describes a situation in which a calibration material behaves such that the equivalence of measurement results between the calibrator and human-tissue-derived materials cannot be achieved. Commutability failures are sometimes the basis for lack of equivalence, but the concept of commutability and the generic comparability and the metrologically defined comparability need to be distinguished from commutability. Definitions from the VIM are given below:

– *Commutability*: commutability of a reference material—property of a reference material, demonstrated by the closeness of agreement between the relation among the measurement results for a stated quantity in this material, obtained according to two given measurement procedures, and the relation obtained among the measurement results for other specified materials [5]. Lack of commutability may be the consequence of matrix effects, interactants, and interferents.

(ix) **Systems of units** are of two types; arbitrary and fundamental. Arbitrary units are the default units when the relationship between the measured phenomenon and established phenomena are unknown or insufficiently defined or definable. When a phenomenon is sufficiently well described and understood, the replacement of arbitrary units with fundamental units enables and promotes the achievement of *traceability* and *comparability* (see below). The internationally adopted fundamental units are those of the International System of Units (SI) [A7].

– *SI units* are based on a limited number of fundamental units with ''fixed'' values than can be assumed to be invariant within specified uncertainty limits. The goal of relating function to structure, a long-standing endeavor of biochemistry, cannot be achieved unless the unit of entity amount is the mole. Although this requirement for traceability to the mole is implicitly recognized, the continued use of arbitrary units, of which there are many for biological entities and activity, greatly impedes linking gene structure to protein structure and protein structure to biological function in any other than a qualitative or descriptive way.
– *Arbitrary units*: at the first or early stages in the discovery of a biological activity, there is frequently no alternative to representing the activity and the entity that possesses the activity by a pool of tissue, e.g., blood plasma that is known to contain the activity. While arbitrary, such a material provides both the reference material that defines the activity and the calibrator that is used in the assay of it. The material is most frequently of biological origin, of uncharacterized homogeneity, and the amount and influence of other molecular entities in the material are uncharacterized. Such complex mixture reference materials are susceptible to ''matrix effects'' that go unrecognized and change upon replacement of the original reference material with replacement batches (see commutability, above).

(x) **Procedure-defined measurand** [6]: a measurand for which the chemical specification included in *c* and the property *f* in the proposed equation are combined. Specifications for the conditions and concentrations of reactants and effectors under which the measurement procedure is performed are comprehensively determined to obtain comparability and traceability of the measured value. Commonly used for the measurement of enzymes and described rigorously as ''catalytic activity concentration of the enzyme as measured by the conversion rate of an indicator entity in a specified system according to a given measurement procedure.'' The procedure-defined measurand is now an ''operationally defined measurand'' [5].

(xi) **Inherent activity**, the term coined here possesses several characteristics previously identified with specific

activity, potency, intrinsic activity, and catalytic efficiency of an enzyme, but in contrast to the previously used terms, restricts the entity amount to an SI-based unit and explicitly separates entity from a property of the entity. Inherent is used to indicate the link to gene (exon) structure. The relation to the term potency is evident from the currently used definitions, with caveats used in pharmacology. Potency as applied reverses the independent and dependent variables when compared to the definition of biological activity proposed here. Potency may be synonymous with specific activity (quotient of a measured value of the biological activity divided by the amount of the entity), although not necessarily so. The term intrinsic activity is not usable because of its existing IUPAC definition.

– *Intrinsic activity* [A1]: intrinsic activity is the maximal stimulatory response induced by a compound in relation to that of a given reference compound. This term has evolved with common usage. It was introduced by Ariëns as a proportionality factor between tissue response and receptor occupancy. The numerical value of intrinsic activity (alpha) could range from unity (for full agonists, i.e., agonist inducing the tissue maximal response) to zero (for antagonists), the fractional values within this range denoting partial agonists.

(xii) ***Chemical species*** [A1]: an ensemble of chemically identical molecular entities that can explore the same set of molecular energy levels on the timescale of the experiment. It is this definition that relates the ''idealized'' reference macromolecular entity to an established IUPAC definition.

(xiii) **Uncertainty** has commonly been associated with the precision of measurement (typically, only under repeatability conditions), an incomplete and inadequate indicator of the quality of a measurement result. Although far beyond the scope of this document, guidance on estimating the uncertainty of a measured value is available from several sources [A10, A11]. The definition proposed here attempts to address uncertainty, but from the perspective of sources of uncertainty, particularly bias, rather than statistical methods for estimating a value for the uncertainty.

– *Uncertainty of measurement*, *measurement uncertainty*, *uncertainty of measurement*, *uncertainty* [5, A8, A10]: parameter, associated with the result of a measurement, that characterizes the dispersion of the values that could reasonably be attributed to the measurand.

The uncertainty of the result of a measurement reflects the lack of exact knowledge of the value of the measurand. The result of a measurement after correction for recognized systematic effects is still only an *estimate* of the value of the measurand because of the uncertainty arising from random effects and from imperfect correction of the result for systematic effects.

In practice, there are many possible sources of uncertainty in a measurement, including: incomplete definition of the measurand; imperfect realization of the definition of the measurand; non-representative sampling—the sample measured may not represent the defined measurand; inadequate knowledge of the effects of environmental conditions on the measurement or imperfect measurement of environmental conditions; personal bias in reading analog instruments; finite instrument resolution; inexact values of measurement standards and reference materials; inexact values of constants and other parameters obtained from external sources and used in the data-reduction algorithm; approximations and assumptions incorporated in the measurement method and procedure; variations in repeated observations of the measurand under apparently identical conditions.

Sources

A1 IUPAC (1997) Compendium of chemical terminology 2nd edn, http://www.iupac.org
A2 Boston University School of Medicine, Department of Pharmacology and Experimental Therapeutics, http://www.bumc.bu.edu/Dept/Content.aspx?DepartmentID=65&PageID=7797
A3 Bliss CI (1957) Some principles of bioassay. American Scientist 45:449–466
A4 Wikipedia®, As of March 1, 2007 http://www.en.wikipedia.org/wiki/Extensibility
A5 United States Pharmacopeia—glossary from the USP guidelines for submitting a request for revision of the USP/NF. Home page at http://www.usp.org
A6 Concise Oxford Dictionary 10th edn (2001) Oxford University Press
A7 Organisation Intergouvernementale de la Convention du Mètre (2006) The International System of Units (SI), International Bureau of weights and measures, BIPM, Paris, France
A8 International vocabulary of basic and general terms in metrology (VIM) 2nd edn (1994)
A9 ICH Q5E (International Conference on Harmonisation) Comparability of biotechnological/biological products subject to changes in their manufacturing process, US FDA
A10 ISO GUM (1993) Guide to the expression of uncertainty in measurement. ISO, Geneva, Switzerland
A11 Eurachem/CITAC (2000) Quantifying uncertainty in analytical measurement, 2nd edn, http://www.measurementuncertainty.org/mu/QUAM2000-1.pdf

References

1. World Health Organization (WHO) (2000) WHO consultation on international biological standards for in vitro diagnostic procedures. WHO, September 2000, Geneva, Switzerland
2. Wyman J (1965) The binding potential, a neglected linkage concept. J Mol Biol 11:631–644
3. Wyman J (1984) Linkage graphs: a study in the thermodynamics of macromolecules. Q Rev Biophys 17(4):453–488
4. Eurachem/CITAC (2003) Guide: traceability in chemical measurement. Available online at http://www.eurachem.ul.pt/guides/Traceab.htm
5. International vocabulary of metrology—basic and general concepts and associated terms (VIM) 3rd edn. Final draft, August 2006
6. Siekmann L, Bonora R, Burtis CA, Ceriotti F, Clerc-Renaud P, Férard G, Ferrero CA, Forest J-C, Franck PFH, Gella F-J, Hoelzel W, Jørgensen PJ, Kanno T, Kessner A, Klauke R, Kristiansen N, Lessinger JM, Linsinger TPJ, Misaki H, Mueller MM, Panteghini M, Pauwels J, Schiele F, Schimmel HG, Vialle A, Weidemann G, Schumann G (2002) IFCC primary reference procedures for the measurement of catalytic activity concentrations of enzymes at 37 degrees C. Part 1. The concept of reference procedures for the measurement of catalytic activity concentrations of enzymes. Clin Chem Lab Med 40(6):631–634
7. Polya G (1962) Mathematical discovery: on understanding, learning and teaching problem solving, vols I, II. Wiley, New York
8. Klotz IM (1964) Chemical thermodynamics: basic theory and methods. Review edition. WA Benjamin, New York
9. Tinoco I, Sauer K, Wang JC (1978) Physical chemistry: principles and applications in biological sciences. Prentice-Hall, Englewood Cliffs, New Jersey
10. Fersht AR (1999) Structure and mechanism in protein science. Freeman, New York
11. The Lewin Group (2005) The value of diagnostics: innovation, adoption and diffusion into healthcare. AdvaMed, Washington, DC. Available online at http://www.advamed.org/publicdocs/july2005hillbriefing.shtml
12. Popper KR (1959) The logic of scientific discovery. Hutchinson, London, UK