Pure App. Chem., Vol. 76, No. 9, pp. 1799–1807, 2004.

INTERNATIONAL FEDERATION OF CLINICAL CHEMISTRY AND LABORATORY MEDICINE SCIENTIFIC DIVISION

COMMITTEE ON MOLECULAR BIOLOGY TECHNIQUES IN CLINICAL CHEMISTRY (C-MBT)

COMMITTEE ON NOMENCLATURE, PROPERTIES AND UNITS (C-NPU)[#]

and

INTERNATIONAL UNION OF PURE AND APPLIED CHEMISTRY CHEMISTRY AND HUMAN HEALTH DIVISION*

PROPERTIES AND UNITS IN THE CLINICAL LABORATORY SCIENCES

PART XVIII. PROPERTIES AND UNITS IN CLINICAL MOLECULAR BIOLOGY

(IUPAC Technical Report)

Prepared for publication by
PEDRO SOARES DE ARAUJO[1,‡], BIANCA ZINGALES[1], PEDRO ALÍA-RAMOS[2], AURORA BLANCO-FONT[3], XAVIER FUENTES-ARDERIU[2], CHRISTINE MANNHALTER[4], KIM VARMING[5], STIG BOJESEN[6] , IVAN BRUUNSHUUS[7], AND HENRIK OLESEN[7]

[1]Department of Biochemistry, Institute of Chemistry, University of São Paulo, Brazil; [2]Servei de Bioquímica Clínica, Ciutat Sanitària i Universitària de Bellvitge, L'Hospitalet de Llobregat, Barcelona, Spain; [3]Institut de Recerca Oncològica, L'Hospitalet de Llobregat, Barcelona, Spain; [4]Department of Laboratory Medicine, University of Vienna Medical School, Vienna, Austria; [5]Department of Clinical Immunology, Aalborg University Hospital, Aalborg, Dennmark; [6]Department of Clinical Biochemistry, Herlev University Hospital, Herlev, Denmark; [7]The National Board of Health, Copenhagen S, Denmark

[#]The combined membership of the IFCC Committee during the preparation of this report (1999–2002) was as follows:
***Chairman*:** U. Forsum (Sweden); ***Members*:** R. Dybkær (Denmark, 1996–); X. Fuentes-Arderiu (Spain, 1992–1999); A. Jabor (Czech Republic, 1998–); W. R. Külpmann (Germany, 1998–); G. Nordin (Sweden, 2000–); P. Soares de Araujo (Brazil, 1994–).

*The combined membership of the IUPAC Task Group during the preparation of this report (2001–2002) was as follows:
***Chairman*:** P. Soares de Araujo (Brazil); ***Members*:** R. Dybkær (Denmark); U. Forsum (Sweden); A. Jabor (Czech Republic); W. Kuelpmann (Germany); G. Nordin (Sweden).

‡Corresponding author: P. S. de Araujo, Depto. de Bioquímica, Instituto de Química, Universidade de São Paulo, Av. Professor Lineu Prestes 748, Bloco 10T, 05508-900 São Paulo, SP, Brazil; Fax: +55 11 3818-5579; E-mail: psdarauj@usp.br

Properties and units in the clinical laboratory sciences
Part XVIII. Properties and units in clinical molecular biology

(IUPAC Technical Report)

Abstract: This document describes the introduction of the concept of property in the field of molecular biology for the presentation of results of clinical laboratory investigations of genes and mutations. It follows the IUPAC–IFCC systematic rules and attempts to create a common base for communication between the clinical laboratory sciences, the medical practitioner, and the molecular biology areas of knowledge.

Because of the plethora of possible structural variations in the outcome of analysis, the designation of components is restricted to the symbols of genes as identified by the Human Genome Nomenclature Database (HUGO). The listing of properties having DNA as system comprises all symbols in the HUGO database except for symbols related to mitochondrial genes, while properties having RNA as system are included only when known to be in actual use. The detailed and accurate presentation of results is made by the laboratory performing the study, and it is recommended that the explicit guidelines given by Antonarakis and den Dunnen be adhered to.

For electronic communication, a code (NPU) is allocated to each property identified. The complete list of properties, which is an integral part of this report, can be accessed at <http://www.iupac.org/publications/pac/2004/7609/7609x1799.html>.

PREFACE

The present document is the 18th part of a series on properties and units in the clinical and environmental human toxicology laboratory sciences initiated in 1987.

The series currently comprises:

I. Syntax and semantic rules [3]
II. Kinds-of-property [4]
III. Elements (of properties) and their code values [5]
IV. Properties and their code values [6]
V. Properties and units in thrombosis and haemostasis [7]
VI. Properties and units in IOC-prohibited drugs [8]
VIII. Properties and units in clinical microbiology [9]
IX. Properties and units in trace elements [10]
X. Properties and units in general clinical chemistry [11]
XI. Coding systems: structure and guidelines [12]
XII. Properties and units in clinical pharmacology and toxicology [13]
XIII. Properties and units in reproduction and fertility [14]
XVI. Properties and units in clinical allergology [15]
XVII. Properties and units for urinary calculi (in preparation)
XVIII. Properties and units in clinical molecular biology (this report)
XIX. Properties and units for transfusion medicine and immunohematology [16]

FOREWORD AND SCOPE

Basic research in biology and medicine and innovations in laboratory methodology have greatly increased the range of properties available to medical practitioners to help them in diagnosis, treatment, and prevention of disease.

The increase in variety of properties examined is now such that the individual physician has insight into or understanding of only a limited number of properties offered to him from the various clinical laboratory specialties.

In the laboratory, local terms (jargon) may be well understood among colleagues, but they are not appropriate for communication with the outside world. Likewise, a laboratory and its local community of users, such as hospital or community physicians, may use a "local dialect" of the language of clinical laboratory sciences that is well understood by all concerned, but when the communication possibilities are wider, even transnational, risks of serious misunderstanding arise.

It is, therefore, essential to promote clear, unambiguous, meaningful, and fully informative communication. Coherence of statements made within and between medical specialties and uniformity in structure of presentation is an objective to be actively pursued. This will facilitate the transfer of information across socio-linguistic barriers.

The purpose of this document is to apply the IUPAC–IFCC recommended syntax structures for request and report, providing formats and names of properties observed in the domain of medical molecular biology, and to facilitate unequivocal written or electronic communication between health care professionals.

For identification of genes, the "approved names" given in the HUGO [1], in the form of symbols, have been used except for symbols of mitochondrial genes. Variations in mitochondrial genes and chromosomes are not dealt with in this technical report. The use of symbols rather than names is in contrast to previous reports for other medical domains where names have been applied systematically. This is because the names of genes are often very extensive and of limited value to the non-specialist. Admittedly, the symbols also are of limited direct informative value, but they are gaining increasing general application and hence meaningful connotation.

One of the main points of this paper is the recognition that any report of a variation of a nucleic acid sequence should include the identity of the sequence referred to. For this purpose, reference is given to "Nomenclature for the description of sequence variation" elaborated by Prof. Stylianos Antonarakis and Dr. Johan T. den Dunnen [2].

The list of properties shown in this document comprises the list of symbols for components as given in the HUGO as of 24 September 2003. This list, which is an integral part of this technical report, can be accessed at <http://www.iupac.org/publications/pac/2004/7609/7609x1799.html>. Since the list of properties contains more than 16 000 entries, it is not suitable for printing on paper. The online version contains links to databases that allow the reader to obtain information about the selected genes. Frequently updated versions of the listing can be obtained at <http://dior.imt.liu.se/cnpu>, the official C-NPU site, also from <http://www.labinfo.dk/English/download_uk.asp> the listing can be downloaded both as an Excel file or as HTML files.

SYNTAX

System(specification)—Component(specification); kind-of-property(specification) = accurate, unambiguous description of findings [17]

CONCEPTS, TERMS, AND DEFINITIONS

system: demarcated arrangement of a set of elements and a set of relationships between these elements [17,18]

EXAMPLES: A portion of DNA from a sample of blood. A portion of RNA from a biopsy.

component: definable part of a system [17,18]

EXAMPLE: GJB1 gene (symbol for gap junction protein, beta 1, 32kDa)(connexin 32, Charcot-Marie-Tooth neuropathy, X-linked)

kind-of-property: common defining aspect of mutually comparable properties

NOTE 1: There is presently no officially approved definition. This definition is for use in this document only.

NOTE 2: In ENV 1614 [18], the term "property" (in a general sense) is used as synonym for kind-of-property.

NOTE 3: A kind-of-property may be qualified by each user to nominal scale, ordinal scale, differential scale, or rational scale. The last three types of kind-of-property are also called kind-of-quantity [17].

nominal scale: scale with a set of possible values for a given kind-of-property that are each a word or symbol without any relation to magnitude [17]

NOTE: The values may be listed in any order according to practical considerations and convention.

designation of property: set of data elements comprising information on system, component, and kind-of-property, and their adherent specifications.

NOTE 1: There is presently no officially approved definition. This definition is for use in this document only.

NOTE 2: Information about identification of patient, time and result is not considered in the entries of this document.

EXAMPLE: DNA(spec.)—A1BG gene; seq.var.

international coding scheme identifier (ICSI): identifier assigned to identify uniquely a registered coding scheme for use in information interchange [19]

EXAMPLE: "NPU" for codes allocated by the C-NPU of the IFCC–IUPAC.

code value: result of applying a coding scheme to an element in a coded set [19]

EXAMPLE: NPU21382 for DNA(spec.)—HLA-F gene; taxon

taxon: kind-of-property indicating non-numerical classification of entities according to values of a set of properties

NOTE 1: There is presently no officially approved definition. This definition is for use in this document only.

EXAMPLE: taxon in "DNA(spec.)—HLA-DQA1 gene; taxon"

sequence variation: kind-of-property indicating change in one or more nucleic acid bases compared to a conventional structure

NOTE 1: There is presently no officially approved definition. This definition is for use in this document only.

NOTE 2: The sequence variation includes such changes as substitution, deletion, insertion, fusion, duplication. Which type is relevant in a given case should be indicated in the result.

NOTE 3: If no sequence variation is found the result should be 0, or alternatively a full description of the outcome should be given.

NOTE 4: The intentional definition is based on the extensional definition given by Antonorakis and Dunnen [2].

STANDARDIZED REQUEST AND REPORT OF CLINICAL LABORATORY RESULTS

The parts of a request and a report are presented in Table 1.

Essential for a request (Table 1) are parts 1 and 2, covering information on patient identification, time or time interval for sampling, and information on the property requested.

The laboratory *report* comprises the three subdivisions 1, 2, and 3.

To each element in part 2 may be added a specification as a parenthetic suffix for clarification and to avoid ambiguity.

Remarks (part 4) with details on findings, interpretation of results, or other.

Thus, the elements of the designation of a property comprise:

System(specification)—Component(specification); kind-of-property(specification)

This is as recommended by IFCC and IUPAC [17] and by the European standard ENV1614:1995 [18].

Table 1 Standard systematic description.

1	Identification and time
	1.1 Identification of patient
	1.2 Date and time(s) of sampling
2	Property
	2.1 System (specification)
	2.2 Component (specification)
	2.3 Kind-of-property (specification)
3	Result
	3.1 Equal-to sign
	3.2 Identification of finding in keeping with ref. [2]
4	Remarks

ELEMENTS OF AN ENTRY

In the series of technical reports under the heading "Properties and Units in the Clinical Laboratory Sciences", the general structure of an entry is as below. The terms recommended are given in boldface, that is: The systematic term for the type of property, the unit, and the coding scheme identifier with a code value.

1. **Name of system and parenthetic specification** spelled out in full, and followed by a long dash (em dash)
2. **Alphanumeric chemical prefixes to component name**
3. **Recommended name of component and parenthetic specification** shifted to the left for alphabetical sorting and searching, and followed by a semicolon
4. **Kind-of-property and parenthetic specification**
5. **Unit**
6. Molar mass (M) for conversion from other units
7. Presently recommended calibrator

8. Previous calibrator(s)
9. Other term(s)
10. Authority: Code value for the international organisation recommending the name of the component or the combined elements of an entry
11. Note with any further information
12. **[NPUXXXXX]. Code value**, intended for interlaboratory transmission between databases
13. Example in abbreviated form

In this technical report, only some of the elements are applicable, as listed below.

1. **Name of system and parenthetic specification** spelled out in full, and followed by a long dash (em dash)
2. **Not applicable**
3. **Recommended name of component and parenthetic specification** shifted to the left for alphabetical sorting and searching, and followed by a semicolon

   Note: In this document, the term for component is substituted by its gene symbol.

4. **Kind-of-property and parenthetic specification**
5. **Not applicable**
6. Not applicable
7. Not applicable
8. Not applicable
9. Other term(s)

   Note: In this document the full term for the component is given in explanation of the symbol used in **3** above.

10. Not applicable
11. Not applicable
12. **[NPUXXXXX]. Code value**, intended for interlaboratory transmission between databases

    Note: The underlined NPU codes contain links to the HUGO database [1] which explain in full the meaning of the gene symbols.

13. Example in abbreviated form

In the examples given, a question mark, "?", has been used to represent a value of a result for a nominal property.

EXAMPLE

1. **DNA(specification.)—**
3. **A1BG gene;**
4. **sequence variation**
9. name of component: alpha-1-B glycoprotein gene
12. **NPU30000**
13. DNA(spec.)—A1BG gene; seq.var. = ?

SYSTEM AND SPECIFICATION

As indicated by the topic, systems are confined to DNA and RNA. Specifications to the system about a parent supersystem, such as "blood", "biopsy from colon cancer", "plasma", "serum", "spinal fluid", "mouth wash", "dried blood spot", are detailed as a note in the EDIFACT or other transmission system.

Component and specification

Components are presented by the symbols for genes from the Human Genome Nomenclature Database or from the HLA Informatics Group [19] with the addition of the word "gene".

Distribution of information between designation of property and result

Detailed information in the designation of properties is omitted, except for a few specified by an OMIM code, as the number of genes precludes granularity beyond the symbol of the gene. If attempted, the number of entries would encompass many millions.

Thus, detailed information according to Antonarakis and Dunnen [2] on findings is given by the laboratory performing the examination and is stated in the form of a result.

Some possible examples on positive and negative findings are presented below. The reference sequence(NM) code values are from ref. [20].

NPU19111 DNA(spec.)—HFE gene; seq.var. = NM_000410.2: [c.187C>G] + [c.187C>G]
NPU19111 DNA(spec.)—HFE gene; seq.var. = no variation at NM_000410.2: [c.187] nor [c.845]
NPU19039 DNA(spec.)—CFTR gene; seq.var. = NM_000492.2: c1522_1524delTTT
NPU19254 DNA(spec.)—ABCC8 gene; seq.var. = NM_000352.2: 4524_4525insCGGCTT

Antonarakis and Dunnen [2] have no suggestions yet as to the presentation of translocations. A possible ad hoc proposal is as:

NPU19599 RNA(Ewing Sarcoma biopsy)—EWSR1 gene; seq.var. = translocation NM_005243.1: 793>ERG NM_004449.2: 695

For properties in molecular biology applied to transfusion medicine and immunohematology, a taxonomic presentation is suggested.

NPU33990 DNA(spec.)—HLA-C gene; taxon = Cw*0102; Cw*0305

REFERENCES

1. Human Genome Nomenclature Database, <http://www.gene.ucl.ac.uk/nomenclature>.
2. S. E. Antonarakis and J. T. den Dunnen. Nomenclature for the description of sequence variation, 7 March 2001; <http://archive.uwcm.ac.uk/uwcm/mg/docs//mut_nom.html>.
3. IUPAC–IFCC (International Union of Pure and Applied Chemistry–International Federation of Clinical Chemistry). Prepared for publication by H. Olesen. *Pure Appl. Chem.* **67**, 1563-1574 (1995); *Eur. J. Clin. Chem. Clin. Biochem*. **33**, 627–636 (1995); *Clin. Chim. Acta* **245**, S5–S21 (1996).
4. IUPAC–IFCC (International Union of Pure and Applied Chemistry-International Federation of Clinical Chemistry). Prepared for publication by D. Kenny and H. Olesen. *Eur. J. Clin. Chem. Clin. Biochem*. **35**, 317–44 (1997).
5. IUPAC–IFCC (International Union of Pure and Applied Chemistry–International Federation of Clinical Chemistry). Prepared for publication by I. Bruunshuus, W. Frederiksen, H. Olesen, I. Ibsen. *Pure Appl. Chem*. **69**, 2577–2582 (1997).
6. IUPAC–IFCC (International Union of Pure and Applied Chemistry–International Federation of Clinical Chemistry). Prepared for publication by H. Olesen, D. Kenny, I. Bruunshuus, I. Ibsen K. Jørgensen, R. Dybkær, X. Fuentes-Arderiu, G. Hill, P. S. de Araujo, C. McDonald. *Pure Appl. Chem*. **69**, 2583–2591 (1997).

7. IUPAC–IFCC (International Union of Pure and Applied Chemistry–International Federation of Clinical Chemistry). Prepared for publication by M. Blombäck, R. Dybkær, K. Jørgensen, H. Olesen, S. Thorsen. *Pure Appl. Chem.* **69**, 1043–1079 (1997); *Eur. J. Clin. Chem. Clin. Biochem*. **33**, 637–660 (1995); *Clin. Chim. Acta* **245**, S23–S28 (1996); *Haemostasis* **71**, 375–394 (1994).
8. IUPAC–IFCC (International Union of Pure and Applied Chemistry–International Federation of Clinical Chemistry). Prepared for publication by H. Olesen, D. Cowan, I. Bruunshuus, K. Klempel, G. Hill. *Pure Appl. Chem*. **69**, 1081–1136 (1997); *Eur. J. Clin. Chem. Clin. Biochem*. **35**, 805–831 (1997); *J. Chromatogr. B* **687**, 157–182 (1996).
9. IUPAC–IFCC (International Union of Pure and Applied Chemistry–International Federation of Clinical Chemistry). Prepared for publication by U. Forsum, H. Olesen, W. Frederiksen, B. Persson. *Pure Appl. Chem*. **72**, 555–745 (2000).
10. IUPAC–IFCC (International Union of Pure and Applied Chemistry–International Federation of Clinical Chemistry). Prepared for publication by R. Cornelis, X. Fuentes-Arderiu, I. Bruunshuus, D. M. Templeton. *Pure Appl. Chem*. **69**, 2593–2606 (1997); *Eur. J. Clin. Chem. Clin. Biochem*. **35**, 833–843 (1997).
11. IUPAC–IFCC (International Union of Pure and Applied Chemistry–International Federation of Clinical Chemistry). Prepared for publication by H. Olesen, I. Ibsen, I. Bruunshuus, D. Kenny, R. Dybkaer, X. Fuentes-Arderiu, G. Hill, P. Soares de Araujo, C. McDonald. *Pure Appl. Chem*. **72**, 747–972 (2000).
12. IUPAC–IFCC (International Union of Pure and Applied Chemistry–International Federation of Clinical Chemistry). Prepared for publication by H. Olesen, D. Kenny, R. Dybkær, I. Ibsen, I. Bruunshuus, X. Fuentes-Arderiu, G. Hill, P. Soares de Araujo, C. McDonald. *Pure Appl. Chem*. **35**, 317–344 (1997).
13. IUPAC–IFCC (International Union of Pure and Applied Chemistry–International Federation of Clinical Chemistry). Prepared for publication by H. Olesen, D. Cowan, R. De la Torre, I. Bruunshuus, M. Rohde, D. Kenny. *Pure Appl. Chem.* **72**, 479–552 (2000).
14. IUPAC–IFCC (International Union of Pure and Applied Chemistry–International Federation of Clinical Chemistry). Prepared for publication by H. Olesen , A. Giwercman, D. M. de Kretser, D. Mortimer, H. Oshima, P. Troen. *Pure Appl. Chem*. **69**, 2621–2628 (1997); *Clin. Chem. Lab. Med*. **36**, 57–65 (1998); *Clin. Chim. Acta* 271, S5–S26 (1998).
15. IUPAC–IFCC (International Union of Pure and Applied Chemistry–International Federation of Clinical Chemistry). Prepared for publication by I. Bruunshuus, L. K. Poulsen, H. Olesen. *Pure Appl. Chem*. **72**, 1067–1205 (2000).
16. IUPAC–IFCC (International Union of Pure and Applied Chemistry–International Federation of Clinical Chemistry). Prepared for publication by K. Varming, U. Forsum, I. Brunshuus, H. Olesen. *Pure Appl. Chem*. **75**, 1477–1600 (2003).
17. IUPAC–IFCC (International Union of Pure and Applied Chemistry–International Federation of Clinical Chemistry). *Compendium of Terminology and Nomenclature of Properties in Clinical Laboratory Sciences* (the "Silver" book), J. C. Rigg, S. S. Brown, R. Dybkær, H. Olesen (Eds.), Blackwell Science, Oxford (1995).
18. CEN/TC 251, 1995. European standard ENV 1614:1995. *Health Care Informatics. Structure for Nomenclature, Classification and Coding of Properties in Clinical Laboratory Sciences*. 5. International Organization for Standardization. International Standard ISO/IEC7826-1:1994 *Information Technology – General Structure for the Interchange of Code Values – Part 1: Identification of Coding Schemes*.
19. Anthony Nolan Trust HLA Informatics Group. <http://www.anthonynolan.org.uk>.
20. National Center for Biotechnology Information LocusLink. <http://www.ncbi.nlm.nih.gov/LocusLink/list.cgi>.

INDEX OF ABBREVIATIONS

HUGO	Human Genome Nomenclature Database
IFCC	International Federation of Clinical Chemistry and Laboratory Medicine
IUPAC	International Union of Pure and Applied Chemistry
MBT	Molecular Biology Techniques in Clinical Chemistry of IFCC
NM	Reference sequence
NPU	Nomenclature, Properties and Units (as International Coding Scheme Identifier)

ACKNOWLEDGMENTS

The authors are indebted to Dr. Joyce Carlson from Clincal Chemistry, University Hospital of Malmö, Sweden for her critical reading of the manuscript.