Pure Appl. Chem., Vol. 85, No. 8, pp. 1725–1758, 2013.
http://dx.doi.org/10.1351/PAC-REC-12-11-23

Glossary of terms used in medicinal chemistry. Part II (IUPAC Recommendations 2013)*

Derek R. Buckle[1,‡], Paul W. Erhardt[2], C. Robin Ganellin[3], Toshi Kobayashi[4], Thomas J. Perun[5], John Proudfoot[6], and Joerg Senn-Bilfinger[7]

[1]DRB Associates, 18 Hillfield Road, Redhill, Surrey, RH1 4AP, UK; [2]University of Toledo, College of Pharmacy, Center for Drug Design and Development, 2801 West Bancroft Street, Toledo, OH 43606-3390, USA; [3]Department of Chemistry, Christopher Ingold Laboratory, University College London, 20 Gordon Street, London, WC1H 0AJ, UK; [4]PhRMA, 4th Floor, Landic II Toranomon Building, 3-7-8 Minato-ku, 105-000 Japan; [5]47731 Old Houston Highway, Hempstead, TX 77445, USA; [6]Boehringer Ingelheim Pharmaceuticals, Inc., 900 Ridgebury Road, P.O. Box 368, Ridgefield, CT 06877, USA; [7]Altana Pharma AG, Byk-Gulden Str. 2, D-78467 Konstanz, Germany

Abstract: The evolution that has taken place in medicinal chemistry practice as a result of major advances in genomics and molecular biology arising from the Human Genome Project has carried with it an extensive additional working vocabulary that has become both integrated and essential terminology for the medicinal chemist. Some of this augmented terminology has been adopted from the many related and interlocked scientific disciplines with which the modern medicinal chemist must be conversant, but many other terms have been introduced to define new concepts and ideas as they have arisen. In this supplementary Glossary, we have attempted to collate and define many of the additional terms that are now considered to be essential components of the medicinal chemist's expanded repertoire.

Keywords: glossary; IUPAC Chemistry and Human Health Division; medicinal chemistry; terminology.

CONTENTS

*Sponsoring body: IUPAC Chemistry and Human Health Division: see more details on p. 1754.

‡Corresponding author: E-mail: drb@drbassoc.co.uk

INTRODUCTION

Since publication of the first "Glossary of Terms used in Medicinal Chemistry" over 10 years ago the practice of medicinal chemistry has undergone a rapid and continuous change. This change, which by necessity has blurred the boundaries between what was traditionally seen as classical medicinal chemistry and its associated scientific disciplines, has resulted in a considerable expansion of related terminology. Medicinal chemists are increasingly required to understand, and interpret, language that was formerly the predominant domain of those involved in a much broader array of biological sciences in which chemistry has an underlying involvement. To reflect this change, the authors have compiled this supplementary Glossary of over 180 additional terms that were not previously collated into a defining document. In compiling this augmented list, a large body of experts actively involved in the practice of medicinal chemistry was consulted, but inevitably with terminology extending into multiple disciplines it is impossible to guarantee that all useful terms have been included.

To avoid a repetition of terms included in the original Glossary we have chosen to keep this supplement as a separate document and to identify it by the designation Part II. By inference, therefore, the earlier Glossary necessarily becomes Part I. Those searching for specific terminology are advised to refer to both Glossaries. For simplicity, it is recommended that Part I [1] is searched prior to Part II, although where descriptive terms defined earlier have been used these are noted as embedded references in the text. Within an entry, when a term is defined elsewhere in the Glossary, it is italicized.

ALPHABETICAL ENTRIES

1. ADMET

Acronym referring to the absorption, distribution, metabolism [1], excretion, and toxicity profile or processes for a xenobiotic [1] upon its administration in vivo.

Note: ADME [1] is also used to delineate these selected parameters within the context of a xenobiotic's pharmacokinetic profile. Because any of the five characteristics may become hurdles during drug development, ADMET behaviour is typically studied and optimized among efficacious analogues during the early drug discovery stage by using in vitro models that attempt to predict such behaviours in clinical studies.

See also *pharmacokinetics*, *drug distribution*.

2. adverse effect
adverse event
adverse drug event

1. (In medicinal chemistry). Undesirable reaction in response to the administration of a drug or test compound.

   Note: In most instances such effects result from off-*target* interactions.

2. (In toxicokinetics). Change in biochemistry, morphology, physiology, growth, development, or lifespan of an organism, which results in impairment of functional capacity or impairment of capacity to compensate for additional stress or increase in susceptibility to other environmental influences.

[2]

3. allosteric antagonist

Compound that binds to a receptor [1] at a site separate from, but actively coupled to, that of the endogenous agonist to reduce actively receptor signals.

Note: The terms "allosteric antagonist" and "*noncompetitive antagonist*" are often synonymous but not necessarily so.

See also *noncompetitive antagonist.*

4. alpha-helix

Secondary 3D structure of a protein or peptide containing a right-handed coiled or spiral conformation, in which every backbone N–H group donates a hydrogen bond to the backbone C=O group of the amino acid four residues earlier.

5. analogue
analog

Chemical compound having structural similarity to a reference compound.

Note: Despite the structural similarity, an analogue may display different chemical and/or biological properties, as is often intentionally the case during design and synthesis to optimize either efficacy or *ADMET* properties within a given series.

See also *analogue-based drug discovery*, *congener*, *follow-on drug.*

6. analogue-based drug discovery
analog-based drug discovery

Strategy for drug [1] discovery and/or optimization in which structural modification of an existing drug provides a new drug with improved chemical and/or biological properties.

Note: Within the context of analogue-based drug discovery, three categories of drug analogues are recognized: Compounds possessing structural, chemical, and pharmacological similarities, termed "direct analogues" and sometimes referred to as "*me-too*" *drugs*; compounds possessing structural and often chemical but not pharmacological similarities, termed "structural analogues"; and structurally different compounds displaying similar pharmacological properties, termed "pharmacological analogues".

See also *ligand-based drug design.*
[3]

7. ATP binding cassette protein
ABC protein

Large gene family of transporter proteins that bind ATP and use the energy to transport substrates (e.g., sugars, amino acids, metal ions, peptides, proteins, and a large number of hydrophobic compounds and metabolites) across lipid membranes.

Note: These proteins have an important role in limiting oral absorption and brain penetration of xenobiotics [1].

See also *efflux pump*, *P-glycoprotein*.
[4]

8. atropisomer

Stereoisomer resulting from hindered rotation about a single bond in which steric hindrance to rotation is sufficient to allow isolation of individual isomers.
[5]

9. attrition rate

Rate of loss of drug [1] candidates during progression through the optimisation and developmental stages while on route to the marketplace.

Note: It has been estimated that for every 10 000 compounds examined during the early stages of biological testing just one reaches the market.

10. autoinduction

Capacity of a drug [1] to induce enzymes that mediate its own metabolism [1].

Note: This often results in lower, often sub-therapeutic, drug exposure on prolonged or multiple dosing.

[6]

11. back-up compound

Molecule selected as a replacement for the lead drug [1] candidate should this subsequently fail during further preclinical evaluation or in clinical studies.

Note: Ideally, a back-up should be pharmacologically equivalent to the lead drug but have significant structural differences. Possession of a distinct core *scaffold* is optimal.

12. best-in-class

Drug acting on a specific *molecular target* that provides the best balance between efficacy [1] and adverse effects.

13. beta-barrel

Secondary 3D protein structure containing a large *beta-sheet* that twists and coils to form a closed structure in which the first-strand hydrogen bonds to the last.

14. beta-sheet

Secondary 3D structure of a protein that takes on a flat, pleated appearance.

Note: Other common protein structural motifs are the *alpha-helix* and *beta-barrel*.

15. beta-sheet breaker

Compound that on binding to a protein disrupts *beta-sheet* formation.

16. bioinformatics

Discipline encompassing the development and utilisation of computational tools to store, analyse, and interpret biological data.

Note: Typically protein or DNA sequence or 3D information.

17. biological agent

Biopolymer-based pharmaceutical, such as a protein, applicable to the prevention, treatment, or cure of diseases or injuries to man.

Note: Biological agents may be any virus, therapeutic serum, toxin, antitoxin, vaccine, blood component or derivative, allergenic product, or analogous products.

[7]

18. biomarker

Indicator signalling an event or condition in a biological system or sample and giving a measure of exposure, effect, or susceptibility.
[2]

19. blockbuster drug

Drug that generates annual sales of USD 1 billion ($\$10^9$) or more.

20. blood–brain barrier (BBB)

Layer of endothelial cells that line the small blood vessels of the brain.

Note 1: These cells form "tight junctions" that restrict the free exchange of substances between the blood and the brain. Such cells are rich in *P-glycoprotein*, which serves to pump substrates back to the peripheral side of the vasculature.

Note 2: Passive diffusion across the BBB is highly dependent on drug lipophilicity, and very few orally active agents acting in the central nervous system have a *polar surface area* greater than 0.9 nm^2.

[8,9]

21. carcinogen

Agent (chemical, physical, or biological) that is capable of increasing the incidence of malignant neoplasms, thus causing cancer.
[10]

22. chemical biology

Application of chemistry to the study of molecular events in biological systems, often using tool compounds.

Note: Distinguished from medicinal chemistry, which is focused on the design and optimisation of compounds for specific *molecular targets*.

[11]

23. chemical database

Specific electronic repository for storage and retrieval of chemical information.

Note 1: Chemical structural information is sometimes stored in string notation such as the *InChI* or *SMILES notations*.

Note 2: Such databases can be searched to retrieve structural information and data on specific or related molecules.

Note 3: A free database of chemical structures of small organic molecules and information on their biological activities is available from PubChem.

[12]

24. chemical diversity

See *diversity*.

25. chemical library
compound library
compound collection

1. Collection of samples (e.g., chemical compounds, natural products, over-expression library of a microbe) available for biological screening.
2. Set of compounds produced through combinatorial chemistry or other means that expands around a single core structure or *scaffold*.

26. chemical space

Set of all possible stable molecules based on a specific chemical entity that interacts at one or more specific *molecular targets*.

27. cheminformatics
chemoinformatics

Use of computational, mathematical, statistical, and information techniques to address chemistry-related problems.
[13]

28. chemogenomics
chemical genomics

Systematic screening of *chemical libraries* of *congeneric* compounds against members of a target family of proteins.
[14]

29. chemokine

Member of a superfamily of proteins with the primary function to control leukocyte activity and trafficking through tissues.
[15]

30. CLOGP values

Calculated 1-octanol/water partition coefficients.

Note: Frequently used in structure–property correlation or quantitative structure–activity relationship [1] studies.

See also *log P*, *log D*.
[2]

31. cluster

Group of compounds that are related by structural, physicochemical, or biological properties.

Note: Organizing a set of compounds into clusters is often used to assess the *diversity* of those compounds, or to develop structure–activity relationship [1] models.

[16–18]

32. co-drug
mutual prodrug [1]

Two chemically linked synergistic drugs [1] designed to improve the drug delivery properties of one or both drugs.

Note: The constituent drugs are indicated for the same disease, but may exert different therapeutic effects via disparate mechanisms of action.

[19]

33. congener

Substance structurally related to another and linked by origin or function.

Note: Congeners may be *analogues* or vice versa but not necessarily. The term "congener", while most often a synonym for "homologue", has become somewhat more diffuse in meaning so that the terms "congener" and "*analogue*" are frequently used interchangeably in the literature.

See also *analogue*, *follow-on drug*.
[2,20,21]

34. constitutive activity

Receptor or enzymatic function displayed in the absence of an agonist [1] or activator.

35. contract research organisation (CRO)

Commercial organisation that can be engaged to undertake specifically defined chemical, biological, safety, or clinical studies.

Note: Typically, such studies are subject to confidentially agreements.

36. covalent drug

Ligand that binds irreversibly to its *molecular target* through the formation of a new chemical bond.

37. cytochrome P450 (CYP450)

Member of a superfamily of heme-containing mono-oxygenases involved in xenobiotic [1] metabolism, cholesterol biosynthesis, and steroidogenesis, in eukaryotic organisms found mainly in the endoplasmic reticulum and inner mitochondrial membrane of cells.
[2]

38. descriptors

See *molecular descriptors*.

39. designed multiple ligands

Compounds conceived and synthesised to act on two or more *molecular targets*.

40. diversity

Unrelatedness of a set of molecules (e.g., building blocks or members of a compound library), as measured by properties such as atom connectivity, physical properties, or computationally generated descriptors.

Note: Inverse of *molecular similarity*.

41. diversity-oriented synthesis (DOS)

Efficient production of a range of structures and *templates* with skeletal and stereochemical diversity as opposed to the synthesis of a specific target molecule.

42. drug cocktail

1. (In drug therapy). Administration of two or more distinct pharmacological agents to achieve a combination of their individual effects.

   Note 1: The combined effect may be additive, synergistic, or designed to reduce side-effects.

   Note 2: This term is often used synonymously with that of "drug combination" but is preferred in order to avoid confusion with medications in which different drugs are included in a single formulation.
2. (In drug testing). Administration of two or more distinct compounds to test simultaneously their individual behaviours (e.g., pharmacological effects in high-throughput screens or drug metabolism [1]).

43. drug combination

See *drug cocktail*.

44. drug delivery

Process by which a drug [1] is administered to its intended recipient.

Note: Examples include administration orally, intravenously, or by inhalation.
See also *drug distribution*, *targeted drug delivery*.

45. drug distribution

Measured amounts of an administered compound in various parts of the organism to which it is given. See also *drug delivery*.

46. druggable target

See *druggability.*

47. druggability

Capacity of a molecular target to be modulated in a favourable manner by a small-molecule drug [1].

Note: It is estimated that only around 10 % of the human genome affords druggable targets. [22,23]

48. drug-like(ness)

Physical and chemical properties in a small molecule that make it likely to perform efficiently as a drug [1].

49. drug repurposing
drug repositioning
drug reprofiling

Strategy that seeks to discover new applications for an existing drug [1] that were not previously referenced and not currently prescribed or investigated.

Note: Various additional synonymous terms have been used to describe the process of drug repurposing. All appear to be used interchangeably.

[24]

50. drug safety

Assessment of the nontolerable biological effects of a drug [1].

Note: Because the nontolerable effects of a drug are directly related to its concentration or dose, safety is generally ranked relative to the dose required to obtain the desirable effect.

See also *therapeutic index*.

51. dual binding site

Presence of two distinct *ligand* binding sites on the same *molecular target*.

52. effective concentration (EC)

Concentration of a substance that produces a defined magnitude of response in a given system.

Note 1: EC_{50} is the median dose that causes 50 % of the maximal response.

Note 2: The term usually refers to an agonist [1] in a receptor system effect and could represent either an increase or a decrease in a biological function.

See also IC_{50}, *effective dose.*
[10,25]

53. effective dose (ED)

Dose of a substance that causes a defined magnitude of response in a given system.

Note: ED_{50} is the median dose that causes 50 % of the maximal response.
See also IC_{50}, *effective concentration.*
[10]

54. efflux pump

Transporter protein located in the membrane of cells that utilizes active transport [1] to move a compound from the internal to the external environment.
See also *P-glycoprotein*, *ATP binding cassette*.

55. epigenetic(s)

Phenotypic change(s) in an organism brought about by alteration in the expression of genetic information without any change in the genomic sequence itself.

Note: Common examples include changes in nucleotide base methylation and changes in histone acetylation. Changes of this type may become heritable.

[2,26]

56. equilibrium solubility

Analytical composition of a mixture or solution that is saturated with one of the components in the designated mixture or solution.

Note 1: Solubility may be expressed in any units corresponding to quantities that denote relative composition, such as mass, amount concentration, molality, etc.

Note 2: The mixture or solution may involve any physical state: solid, liquid, gas, vapour, supercritical fluid.

Note 3: The term "solubility" is also often used in a more general sense to refer to processes and phenomena related to dissolution.

See also *intrinsic solubility*, *kinetic solubility*, *solubility*, *supersaturated solution*.
[27–29]

57. fast follower

Compound selected as a rapid successor to a lead drug [1] candidate.

Note: Fast followers usually possess a marked increase in one or more pharmacological/pharmaceutical characteristics such as potency, efficacy [1], therapeutic index, or physicochemical parameters (e.g., *solubility*).

58. fingerprint

Representation of a compound or *chemical library* by attributes (descriptors) such as atom connectivities, 3D structure, or physical properties.
[30,31]

59. first-in-class

First drug acting on a hitherto unaddressed molecular target to reach the market.

60. follow-on drug

Drug [1] having a similar mechanism of action to an existing drug.

Note: Compounds may be of the same or different chemical class. A therapeutic advantage over *first-in-class* drugs must be demonstrated for regulatory approval.

See also *analogue*, *analogue-based drug design*, *congener.*
[32,33]

61. fragment

Low-molar-mass *ligand* (typically smaller than 200 Da) that binds to a target with low affinity [1] but high *ligand efficiency*.

Note: Typically fragments have affinities in a concentration interval from 0.1 to 1.0 mM.

62. fragment-based lead discovery

Screening libraries of low-molar-mass compounds (typically 120–250 Da) using sensitive biophysical techniques capable of detecting weakly binding lead compounds.

Note: X-ray structures are frequently used to drive the optimisation of fragment *hits* to *leads*.

See also *fragment*, *ligand efficiency.*
[34]

63. frequent hitter

Structural feature that regularly results in a positive response in a variety of *high-throughput* or primary screens.

Note: Such compounds often exert their actions through nonspecific mechanisms and are therefore unreliable leads.

[35]

64. genomics

Science of using DNA- and RNA-based technologies to demonstrate alterations in gene expression.
[2]

65. good laboratory practice (GLP)

Set of principles that provides a framework within which laboratory studies are planned, performed, monitored, recorded, reported, and archived.

Note: These studies are undertaken to generate data by which the hazards and risks to users, consumers, and third parties, including the environment, can be assessed for pharmaceuticals (only preclinical studies), agrochemicals, cosmetics, food additives, feed additives and contaminants, novel foods, biocides, detergents, etc. GLP helps assure regulatory authorities that the data submitted are a true reflection of the results obtained during the study and can therefore be relied upon when making risk/safety assessments.

[36]

66. good manufacturing practice (GMP)

Quality assurance process that ensures that medicinal products are consistently produced and controlled to the standards appropriate to their intended use.

Note: Quality standards are those required under marketing authorisation or product specification. GMP is concerned with both production and quality control.

[37]

67. G-protein
guanine binding protein

Member of a family of membrane-associated proteins which on activation by cellular receptors [1] lead to signal transduction.
See also *G-protein-coupled receptor*.

68. G-protein-coupled receptor (GPCR)

Large family of cell surface receptors [1] in which seven portions of the protein cross the cellular membrane and are linked to internal *G-proteins*.

Note 1: Interaction of these receptors with extracellular ligands activates signal transduction pathways and, ultimately, cellular responses.

Note 2: G-protein-coupled receptors are found only in eukaryotes.

See also *G-protein.*

69. green chemistry

Invention, design, and application of chemical products and processes to reduce or eliminate the use and generation of hazardous substances.
[38]

70. high-throughput screening (HTS)

Method for the rapid assessment of the activity of samples from large compound collections.

Note 1: Typically, these assays are carried out in microplates of at least 96 wells using automated or robotic techniques.

Note 2: The rate of at least 10^5 assays per day has been termed "ultra-high-throughput screening" (UHTS).

[25]

71. hit

Molecule that produces reproducible activity above a defined threshold in a biological assay and whose structural identity has been established.

Note: Hits typically derive from *high-throughput screening* initiatives or other relatively extensive primary assays and do not become true hits until fully validated.

72. hit expansion

Generation of additional compound sets that contain chemical motifs and *scaffolds* that have activity in the primary screen.

Note: This methodology permits the identification of additional *hits* and new *scaffolds* and develops structure–activity relationships [1] around existing hits.

73. hit-to-lead chemistry

Process by which a proven molecule or series derived from *high-throughput screening* or primary screens is chemically optimised to a viable *lead* or series.

74. homology model

Computational representation of a protein built from the 3D structure of a similar protein or proteins using alignment techniques and homology arguments.

75. hydrophobic fragmental constant

Representation of the lipophilicity [1] contribution of a constituent part of a structure to the total lipophilicity.

[39]

76. hydrophobic interaction

Entropically driven favourable interaction between nonpolar substructures or surfaces in aqueous solution.

See [2] for alternative definition.

77. IC_{50} (inhibitory concentration 50)

The concentration of an enzyme inhibitor or receptor antagonist [1] that reduces the enzyme activity or agonist response by 50 %.

Note: IC_{50} values are influenced by experimental conditions (e.g., substrate or agonist [1] concentration, which should be specified).

Related terms: *inhibition constant*, K_i

[25]

78. inhibition constant, K_i

1. Equilibrium dissociation constant of an enzyme-inhibitor complex: $K_i = [E][I]/[EI]$.
2. The equilibrium dissociation constant of a receptor-ligand complex.

Note: This value is usually obtained through competition binding experiments, where the K_i is determined after the IC_{50} obtained in a competition assay performed in the presence of a known concentration of labeled reference ligand.

Related term: IC_{50}
[25]

79. in silico screening

See *virtual screening*.

80. International Chemical Identifier (InChI)

Non-proprietary identifier for chemical substances that can be used in printed and electronic data sources to enable easier linking of diverse data compilations.

Note: This IUPAC notation frequently replaces the earlier *SMILES* notation.
[40]

81. intercalation

Thermodynamically favourable, reversible inclusion of a molecule (or group) between two other molecules (or groups).
Examples include DNA intercalation.

82. intrinsic solubility

Equilibrium solubility of the uncharged form of an ionisable compound at a pH where it is fully unionised.

$$HA(s) = HA(aq) \quad (1)$$

$$HA(aq) = H^+ + A^- \quad (2)$$

The intrinsic solubility can be determined from the analytical composition at a pH where [HA] is very much greater than $[A^-]$.
See also *equilibrium solubility*, *kinetic solubility*, *solubility*, *supersaturated solution*.
[27–29]

83. investigational new drug (IND)

Drug [1] not yet approved for general use by the national authority, such as the Food and Drug Administration of the United States of America, but undergoing clinical investigation to assess its safety and efficacy.

84. ionotropic receptor

Transmembrane ion channel that opens or closes in response to the binding of a *ligand*.

85. kinase
phosphotransferase

Enzyme [1] that transfers a phosphate group from high-energy donor molecules, such as ATP, to specific target molecules.

86. kinetic solubility
turbidimetric solubility

Composition of a solution with respect to a compound when its induced precipitate first appears.
See also *equilibrium solubility*, *intrinsic solubility*, *solubility*, *supersaturated solution.*
[41]

87. lead

Compound (or compound series) that satisfies predefined minimum criteria for further structure and activity optimization.

Note: Typically, a lead will demonstrate appropriate activity, selectivity, tractable structure–activity relationship [1] and have confirmed activity in a relevant cell-based assay.

See *lead validation.*
[24]

88. lead validation

Process by which a *lead* compound is authenticated by the confirmation of its expected pharmacological properties.

Note: Usually, a cluster of structurally similar compounds showing discernable structure–activity relationships [1] will support the validation process.

89. ligand

Ion or molecule that binds to a *molecular target* to elicit, block, or attenuate a biological response.

90. ligand-based drug design

Method of drug discovery and/or optimization in which the pursuit of new structures and/or structural modifications is based upon one or more *ligands* known to interact with the *molecular target* of interest.

Note: This approach is applicable even when no structural detail of the target is known. In such cases a series of *analogues* is usually prepared and tested to produce structure–activity relationship data [1] that can be extrapolated to indirectly derive a topographical map of the biological surface.

See also *analogue-based drug discovery*, *structure-based drug design*.

91. ligand efficiency (LE)

Measure of the free energy of binding per heavy atom count (i.e., non-hydrogen) of a molecule.

Note 1: It is used to rank the quality of molecules in drug discovery, particularly in *fragment-based lead discovery*.

Note 2: An LE value of 1.25 kJ mol^{-1} (non-hydrogen atom)$^{-1}$ is the minimum requirement of a good *lead* or *fragment*.

[34,42]

92. ligand lipophilic efficiency (LLE)
lipophilic efficiency

Parameter used to identify *ligands* with a high degree of specific interaction towards the desired *molecular target*.

Note 1: The potency of a *ligand* towards a *molecular target* may be dominated by nonspecific partitioning from the aqueous phase. It can be advantageous to separate out the non-specific component of the potency in order to identify more specific interactions; typically using an equation such as: LLE, symbol E_{LL}, is defined by the logarithm of the potency minus a lipophilicity measure, where a typical example would be

$$E_{LL} = -\log (IC_{50}) - \log P$$

Note 2: E_{LL} can be regarded as part of a thermodynamic cycle used as a complementary measure to potency in the search for specific target interactions. In this case, the dissociation constant, K_d, is a more appropriate measure than IC_{50} since it refers to the Gibbs energy of the binding process

$$E_{LL} = -\log K_d - \log P$$

[43]

93. Lipinski's rule

See *rule of five*.

94. log *D*
lg *D*

Logarithm of the apparent partition coefficient at a specified pH.
See *CLOGP*, *log P*.

95. log *P*
lg *P*

Measure of the lipophilicity [1] of a compound by its partition coefficient between an apolar solvent (e.g., 1-octanol) and an aqueous buffer.

Thus, *P* is the quotient of the concentration of non-ionised drug in the solvent divided by the respective concentration in buffer.
See *CLOGP*, *log D*.

96. Markush structure

Generalised formula or description for a related set of chemical compounds used in patent applications and chemical papers.

97. metabotropic receptor

Receptor [1] that modulates electric potential-gated channels via *G-proteins*.

Note: The interaction can occur entirely within the membrane or by the generation of diffusible second messengers. The involvement of G-proteins causes the activation of these receptors to last tens of seconds to minutes, in contrast with the brief effect of *ionotropic receptors*.

[44]

98. me-too drug

Term applied to drug analogues that offer no significant advantage over the prototype compound.

Note: Continued use of this term is not recommended.

See *analogue*, *follow-on drug*.

[21]

99. microarray

Planar surface where assay reagents and samples are distributed as sub-microlitre drops.

Note: This screening format is a direct offshoot of genomic microarray technologies and makes use of ultra-low-volume miniaturization provided by nanodispensing technologies.

[25]

100. microRNA (miRNA)

Small single-stranded RNA molecules that play a significant role in the post-transcriptional regulation of gene expression.

Note: MicroRNA usually comprises approximately 22 nucleotides.

[45]

101. molecular descriptor

Parameter that characterizes a specific structural or physicochemical aspect of a molecule.

[46]

102. molecular diversity

See *diversity*.

103. molecular dynamics

Computational simulation of the motion of atoms in a molecule or of individual atoms or molecules in solids, liquids, and gases, according to Newton's laws of motion.

Note: The forces acting on the atoms, required to simulate their motions, are generally calculated using molecular mechanics force fields.

See [2] for alternative definition.

[47]

104. molecular similarity

Measure of the coincidence or overlap between the structural and physicochemical profiles of compounds.

105. molecular target

Protein (e.g., receptor [1], enzyme [1], or ion channel), RNA, or DNA that is implicated in a clinical disorder or the propagation of any untoward event.

Note: Usually, biochemical, pharmacological, or genomic information supporting the role of such a target in disease will be available.

106. multidrug resistance (MDR)

Characteristic of cells that confers resistance to the effects of several different classes of drugs.

Note: There are several forms of drug resistance. Each is determined by genes that govern how cells will respond to chemical agents. One type of multidrug resistance involves the ability to eject several drugs out of cells (e.g., *efflux pumps* such as *P-glycoprotein*).

[48]

107. multiparameter optimisation (MPO)

Drug-likeness penetrability algorithm derived from *CLOGP*, *c*log*D*, molar mass, *topological polar surface area*, number of hydrogen-bond donors, and pK_a.

Note: The MPO desirability score is larger or equal to 4 on a scale of 0 to 6.

[49]

108. multitarget-directed ligand (MTDL)
multitarget drug

Ligand acting on more than one distinct *molecular target*. Targets may be of the same or different mechanistic classes.

109. multitarget drug discovery (MTDD)

Deliberate design of compounds that act on more than one *molecular target*.

110. Murcko assembly

Core *scaffold* of a molecule that remains after all chain substituents that do not terminate in a ring are removed. Single atoms connected by a double bond are typically also retained.
[50]

111. neglected disease

Term used to emphasise an imbalance between therapeutic needs and resource allocation.

Examples include tropical infections that are especially endemic in regions of Africa, Asia, and Central and South America.

Note 1: Different groups define neglected diseases differently, but typically they include diseases such as schistosomiasis, Chagas disease, tuberculosis, and malaria, as well as other parasitic and vector-borne diseases.

Note 2: The terms "tropical disease" and "neglected disease" are sometimes indistinguishable or used interchangeably.

See also *orphan disease*.

112. neural network

Statistical analysis procedure based on models of nervous system learning in animals.

Note: Neural networks have the ability to "learn" from a collection of examples to discover patterns and trends. These data-mining techniques can be used in forecasting or prediction.

[51]

113. neutral antagonist (in pharmacology)

Ligand that blocks the responses of a receptor [1] to both agonists [1] and inverse agonists [1] with the same intensity. It binds to the receptor without evoking any change of conformation or change to the ratio of activated to inactivated conformations.

Note: Perfect neutral antagonism is difficult to achieve.

114. new chemical entity

Drug [1] that contains no active moiety previously approved for use by the national authority, such as the U.S. Food and Drug Administration.
See also *new molecular entity*.
[52]

115. new molecular entity

Active ingredient that has never before been approved in any form by the national authority, such as the U.S. Food and Drug Administration.
See also *new chemical entity*.
[52]

116. noncompetitive antagonist

Functional antagonist [1] that either binds irreversibly to a receptor [1] or to a site distinct from that of the natural agonist [1].
See *allosteric antagonist*.

117. nuclear hormone receptor
nuclear receptor

Ligand-activated transcription factor that regulates gene expression by interacting with specific DNA sequences upstream of its target gene(s).
[53]

118. obviousness

Term associated with intellectual property wherein the latter's *patentability* is assessed relative to the combination of more than one item of "prior art".

Note: To be patentable within the context of medicinal chemistry, a given compound must be: (i) novel, in that its specific arrangement of atoms has never been previously disclosed; (ii) non-obvious, in that its specific arrangement of atoms is not readily suggested to be of benefit by a person having ordinary skill in the art upon considering two or more other, previously disclosed structures; and, (iii) useful, in that it should have some benefit, the disclosure of the latter encompassing a valid "reduction to practice".

See also *patentability*.

119. off-target effect

Pharmacological action induced by any molecule at molecular/biological sites distinct from that for which it was designed.

Note: Such effects are dose-dependent and may be beneficial, adverse, or neutral.

120. orphan disease

Disease for which drug research, development, and marketing is economically unfavourable.

Note 1: The poor commercial environment could be due to a lack of economic incentives or a lack of understanding of the diseases or a combination of both.

Note 2: Sometimes the term "rare disease" is used synonymously with orphan disease, although there is a slight difference. For example, a rare disease is so uncommon that there is no drug development effort.

Note 3: Which diseases are classified as orphan depends strongly on the country that classifies it. In the United States, for example, any disease affecting less than 200 000 people is considered an orphan or rare disease. Europe and countries such as Japan, Australia, and Singapore have a different definition.

See also *orphan drug*, *neglected disease*.

121. orphan drug

Pharmaceutical agent that has been approved specifically to treat a rare and commercially unfavourable medical condition.

See also *orphan disease*.

122. orphan receptor

Receptor for which an endogenous *ligand* has yet to be identified.

123. parallel synthesis

Simultaneous preparation of sets of discrete compounds in arrays of physically separate reaction vessels or microcompartments without interchange of intermediates during the assembly process.

[54]

124. patentability

Set of criteria that must be satisfied in order to achieve commercial exclusivity for an invention.

Note: These criteria are essentially the same in all major countries and include: suitability, novelty, inventiveness, utility, and the provision of an adequate description.

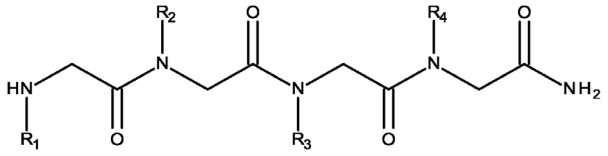
See also *obviousness*.
[55]

125. peptidase

See *proteinase*.

126. peptoid

Peptide-like oligomer consisting of repeating *N*-substituted glycine units



where R_1 to R_4 are typically alkyl or aryl and may be the same or different.
[54]

127. P-glycoprotein (Pgp)

ATP-binding cassette transporter responsible for the efflux of small molecules from cells.

Note: P-glycoproteins can play a major role in limiting brain penetration and restricting the intestinal absorption of drugs. Their over-expression in cancer cells becomes a common mechanism of *multidrug resistance*.

See also *blood–brain barrier*, *efflux pump*.
[56]

128. pharmacogenetics

Study of inherited differences (variation) in drug metabolism [1] and response.
See *pharmacogenomics.*
[57]

129. pharmacogenomics

General study of all of the many different genes that determine drug behaviour.

Note: The distinction between the terms *pharmacogenetics* and *pharmacogenomics* has blurred with time. and they are now frequently used interchangeably.

See *pharmacogenetics*.
[57]

130. phase 0 clinical studies
exploratory investigational new drug

Exploratory first-in-human trials that involve microdosing of drug [1] to allow the assessment of pharmacokinetic parameters with limited drug exposure.

Note: These trials have no therapeutic or diagnostic intent but are designed to assist decision making by providing bioavailability, metabolism, and other limited data from a small number of patients.

See also *phase I*, *II*, *III*, *IV clinical studies.*
[52]

131. phase I clinical studies

Initial introduction of an *investigational new drug* into humans.

Note: These studies are designed to determine the metabolic and pharmacologic actions of the drug [1] in humans, the side-effects associated with increasing doses, and, if possible, to gain early evidence on effectiveness.

See also *phase 0*, *II*, *III*, *IV clinical studies*.
[52]

132. phase II clinical studies

Controlled clinical studies conducted in a limited number of individuals to obtain some preliminary data on the effectiveness of an *investigational new drug* for a particular indication or indications in patients with the disease or condition.

Note: Phase II clinical trials have two subclasses, IIa and IIb. Phase IIa trials are essentially pilot clinical trials designed to evaluate efficacy (and safety) in selected populations of patients with the disease or condition to be treated, diagnosed, or prevented. Phase IIb trials extend those of Phase IIa to well-controlled trials that evaluate the same parameters in similar patient populations.

See also *phase 0*, *I*, *III*, *IV clinical studies.*
[52]

133. phase III clinical studies

Expanded controlled and uncontrolled trials in humans.

Note: These trials are performed after preliminary evidence suggesting effectiveness of the drug [1] has been obtained in *phase II*, and are intended to gather the additional information about effectiveness and safety that is needed to evaluate the overall benefit/risk relationship of the drug.

See also *phase 0*, *I*, *II*, *IV clinical studies.*
[52]

134. phase IV clinical studies

Extended post-marketing studies in humans.

Note: These trials are designed to broaden information concerning treatment risks, benefits, and optimal drug use.

See also *phase 0*, *I*, *II*, *III clinical studies.*
[52]

135. phenotypic screening

Evaluation of compounds (small molecules, peptides, *siRNA*, etc.) in cells, tissues, or oganisms for their ability to modify the system in a measurable manner.

Note: Phenotypic screening differs from target-based screens in that it assesses the overall response of the compound under investigation rather than its specific response on a purified molecular target. Advances in molecular biology resulted in a marked shift away from phenotypic screening, although the latter is now regaining popularity.

[58]

136. phosphotransferase

See *kinase*.

137. pipeline

Discovery and development compound portfolio of a pharmaceutical company or research organization.

138. pivotal study

Experiment that provides strong support for or against the drug [1] or *molecular target* under investigation.
See also *proof of concept*.

139. polar surface area
topological polar surface area

Surface area over all polar atoms (usually oxygen and nitrogen), including any attached hydrogen atoms, of a molecule.

Note: Polar surface area is a commonly used metric (c.f. *molecular descriptor*) for the optimisation of cell permeability. Molecules with a PSA of greater than 1.4 nm^2 are usually poor at permeating cell membranes. For molecules to penetrate the *blood–brain barrier*, the polar surface area should normally be smaller than 0.6 nm^2, although values up to 0.9 nm^2 can be tolerated.

See also *blood–brain barrier*.
[59]

140. polymorphism

Ability of a compound to exist in more than one crystalline form (polymorph) with each having a different arrangement or conformation of the molecules within the crystal lattice.

Note: Polymorphs generally differ in their melting points, *solubility*, and relative intestinal absorption such that optimal polymorphs can markedly enhance the attractiveness of some drugs.

[60]

141. positron emission tomography (PET)

Imaging technique used to visualize small amounts of a compound in biological tissues by the use of radionuclide labels. These radionuclides, such as ^{11}C, ^{18}F, ^{13}N, and ^{15}O, are positron emitters.

142. potential genotoxic impurity

Confirmed or suspected presence of one or more known genotoxic materials in an active pharmaceutical ingredient.

Note: Typically, such impurities arise from reactive intermediates involved in the synthetic pathway and have the potential to adversely affect subsequent genotoxicity tests.

143. preclinical candidate (PCC)
safety assessment candidate

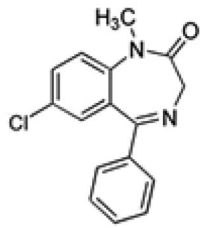
Optimised *lead* compound successfully passing key screening, selectivity, and physicochemical criteria sufficient to warrant further detailed pharmacological and pharmacokinetic evaluation in animal models.

Note: Critical studies usually include bioavailability, therapeutic efficacy in an appropriate disease model, and side-effect profiling.

144. privileged structure

Substructural feature that confers desirable (often drug-like) properties on compounds containing that feature. They often consist of a semi-rigid *scaffold* that presents multiple hydrophobic residues without undergoing hydrophobic collapse.

Note 1: For example, diazepam (below) in which the diphenylmethane moiety prevents association of the aromatic rings.



Note 2: Such structures are commonly found to confer activity against different targets belonging to the same receptor [1] family.

[61]

145. proof of concept (in pharmacology)

Procedure by which a specific therapeutic mechanism or treatment/diagnostic paradigm is shown to be beneficial.

Note: Similar to, and often simultaneously associated with, that for a new drug [1] candidate. This process usually involves an early supporting step prior to clinical testing and final validation within human studies.

See also *lead validation*, *pivotal study*.

146. proteinase
protease

Enzyme [1] that catalyses the hydrolysis of proteins.

Note: Usually, several proteolytic enzymes are necessary for the complete breakdown of polypeptides to their amino acids.

[2]

147. protein data bank (PDB)

Repository for the 3D structural data of large biological molecules including proteins and nucleic acids.

Note: These high-resolution structures, generated predominantly by X-ray or NMR spectroscopic techniques, provide a major resource for structural biology.

[62]

148. protein–protein interaction (PPI)

Association of one protein with one or more other proteins to form either homo- or heteromeric proteins.

Note: Such associations are common in biological systems and are responsible for the regulation of numerous cellular functions in addition to the mediation of disease morphology where aberrant interactions play a significant role.

149. prototype drug

Early compound that has biological properties suitable for *target* validation but may not necessarily be adequate for clinical studies.

150. QTc interval

Time between the start of the Q wave and the end of the T wave in the heart's electrical cycle. When corrected for individual heart rate it becomes known as the corrected QT, or QTc interval.

Note: Significant prolongation of the QTc interval by pharmaceutical agents can induce life-threatening ventricular arrythmia (Torsades de Pointes), typically by interacting with the hERG channel.

[63]

151. retrosynthesis

Process of conceptually deconstructing complex molecules into simpler fragments capable of chemical manipulation to reform the parent compound.

152. RNA interference (RNAi)

Physiological process used by cells to attenuate, or silence, the activity of specific genes.

Note: RNAi offers a means to manipulate gene expression experimentally and to probe gene function on a whole-genome scale.

See also *siRNA*.

153. rule of five

Set of *molecular descriptors* used to assess the potential oral bioavailability of a compound.

Note 1: Characterized by a mass of less than 500 Da, less than or equal to 5 hydrogen bond donors, less than or equal to 10 hydrogen bond acceptors (usually using the N+O count as a surrogate for the number of hydrogen bond acceptors), and a CLOGP less than or equal to 5.

Note 2: Frequently used to profile a *chemical library* or *virtual chemical library* with respect to the proportion of *drug-like* members that it contains. Often used as a surrogate for "drug-likeness".

Note 3: While these criteria are frequently referred to as Lipinski's rules, the de-personalised term "rule of five" is preferred.

[64]

154. rule of three

Set of *molecular descriptors* used to assess the quality of *hit* or *lead* molecules.

Note 1: Most commonly applied to *fragments* that ideally are characterized by a mass of less than 300 Da, less than or equal to 3 hydrogen bond donors, less than or equal to 3 hydrogen bond acceptors, and a CLOGP of less than or equal to 3. In addition, the number of rotatable bonds should average or be less than 3 and the *polar surface area* about 0.6 nm^2.

Note 2: Often used to distinguish "lead-like" from "*drug-like*" molecules.

See *fragment-based lead discovery*.

[65]

155. scaffold
template

Core portion of a molecule common to all members of a *chemical library* or compound series.

156. scaffold hopping

Exchange of one *scaffold* for another while maintaining molecular features that are important for biological properties.

[66]

157. similarity

See *molecular similarity.*

158. siRNA

See *small inhibitory double-stranded RNA*.

159. site-directed mutagenesis

Molecular biology technique in which mutations are created at one or more defined sites in a DNA molecule.

Note: Typically used in *molecular target* validation and to determine whether specific amino acids are involved at *ligand* or substrate interaction sites.

160. small inhibitory double-stranded RNA (siRNA)
small interfering double-stranded RNA

Small RNA fragments that can combine with specific genes to silence their expression.

Note: siRNAs are powerful tools that are often used for manipulating gene expression during *molecular target* validation.

See also *RNAi.*
[67]

161. SMILES (simplified molecular input line entry system) notation

String notation used to describe the atom type and connectivity of molecular structures.

Note: Primarily used to input chemical structures into electronic databases and now frequently replaced by the *InChI* notation.

[68]

162. solubility

Analytical composition of a mixture or solution that is saturated with one of the components of the mixture or solution, expressed in terms of the proportion of the designated component in the designated mixture or solution.

Note 1: Solubility may be expressed in any units corresponding to quantities that denote relative composition, such as mass, amount concentration, molality, etc.

Note 2: The mixture or solution may involve any physical state: solid, liquid, gas, vapour, supercritical fluid.

Note 3: The term "solubility" is also often used in a more general sense to refer to processes and phenomena related to dissolution.

See *equilibrium solubility*, *intrinsic solubility*, *kinetic solubility*, *supersaturated solution.*
[29]

163. spare receptor
receptor reserve

Residual binding site still available to an endogenous *ligand* after sufficient sites have already been filled to elicit the maximal response possible for that particular biological system.

Note: Xenobiotics [1] such as drugs [1] may similarly interact with such receptors [1], but by definition, their identification and quantification occurs via use of the natural *ligand*.

164. stem cell

Multipotent cell with mitotic potential that may serve as a precursor for many kinds of differentiated cells.

Note: Unipotent stem cells can differentiate into one mature cell type only.
[2]

165. structural alert

Chemical features present in a *hit* or *lead* molecule indicative of potential toxicity.

Note: Typically, such features include chemically reactive functionality and components known to metabolise to chemically reactive entities. Examples include anhydrides, aromatic amines, and epoxides.

166. supersaturated solution

Solution that has a greater composition of a solute than one that is in equilibrium with undissolved solute at specified values of temperature and pressure.
See also *equilibrium solubility*, *intrinsic solubility*, *kinetic solubility*.
[29]

167. systems biology

Integration of high-throughput biology measurements with computational models that study the projection of the mechanistic characteristics of metabolic and signalling pathways onto physiological and pathological phenotypes.
[69]

168. target

See *molecular target*.

169. targeted drug delivery

Approach to target a drug to a specific tissue or *molecular target* using a prodrug [1] or antibody recognition systems.
See site-specific delivery in ref. [1].

170. target validation

Process by which a protein, RNA, or DNA is implicated in a biological pathway thought to be of relevance to a disease or adverse pathology.

Note: Typically, validation will involve location of the *molecular target* in relevant cells, organs, or tissues, evidence for its up-regulation/activation in the disorder, and the ability to attenuate adverse responses by agents known to interfere with the *target*.

171. tautomer

Structural isomer that can readily convert to another form that differs only by the attachment position of a hydrogen atom and the location of double bond(s).

Note: In most cases, these isomers are formed by a proton shift to or from heteroatoms such as O, N, or S as typified by the enol and keto forms of carbonyl compounds. Tautomers rapidly interconvert by proton transfer and are usually in equilibrium with one another.

See also *tautomerism*.
[1,70]

172. tautomerism

Reversible interconversion of two different *tautomers*.

Note: For an expanded definition, see [1].

173. template

See *scaffold*.

174. therapeutic index
therapeutic ratio

Ratio of the exposure/concentration of a therapeutic agent that causes beneficial effects to that which causes the first observed adverse effect.

Note: A commonly used measure of therapeutic index is the toxic dose of a drug for 50 % of a population divided by the minimum effective dose for 50 % of a population.

175. topological polar surface area

See *polar surface area*.

176. training set

Specific group of compounds selected for characterisation of both the *molecular descriptors* and the measured values of the targeted property.

Note: Statistical methods applied to the set are used to derive a function between the molecular descriptors and the targeted property.

177. ultra-high-throughput screening (UHTS)

See *high-throughput screening*.

178. unmet medical need

Term used for diseases or other disorders for which no optimal therapeutic options exist.

179. virtual chemical library

Collection of chemical structures constructed solely in electronic form or on paper.

Note: The building blocks required for such a library may not exist, and the chemical steps for such a library may not have been tested. These libraries are used in the design and evaluation of possible libraries.

See *virtual screening*.
[71]

180. virtual screening
in silico screening

Evaluation of compounds using computational methods.

Note: The source of the model could be a macromolecular structure or based on physicochemical parameters or *ligand* structure–activity relationships.

181. volume of distribution (V_d)

Apparent (hypothetical) volume of fluid required to contain the total amount of a substance in the body at the same concentration as that present in the plasma assuming equilibrium has been attained.
[10]

181. wild-type receptor

Receptor that occurs naturally in human and other species.

MEMBERSHIP OF SPONSORING BODIES

Membership of the IUPAC Chemistry and Human Health Division Committee during the preparation of this report (2008–2011) was as follows:

***President*:** D. M. Templeton (Canada, 2008–2011); ***Secretary*:** M. S. Chorghade (USA, 2008–2009); M. Schwenk (Germany, 2010–2011); ***Past President*:** P. W. Erhardt (USA, 2008–2009); ***Vice President*:** F. Pontet (France, 2010–2011); ***Titular Members*:** O. Andersen (Denmark, 2008–2011); S. O. Bachurin (Russia, 2010–2011); D. R. Buckle (UK, 2010–2011); X. Fuentes-Arderiu (Spain, 2008–2011); H. P. A. Illing (UK, 2010–2011); M. N. Liebman (USA, 2008–2009); Y. C. Martin (USA, 2010–2011); T. Nagano (Japan, 2010–2011); M. Nordberg (Sweden, 2008–2009); F. Pontet (France, 2008–2009); F. Sanz (Spain, 2008–2009); G. Tarzia (Italy, 2008–2010); ***Affiliate Members*:** C. R. Ganellin (UK, 2008–2011); T. J. Perun (USA, 2008–2011); J. H. Duffus (UK, 2008–2011).

Active membership of the Subcommittee on Medicinal Chemistry and Drug Development (2008–2011) was as follows: C. R. Ganellin (UK, *Chair*); J. Proudfoot (USA, *Secretary*); S. O. Bachurin (Russia); E. Breuer (Israel); D. R. Buckle (UK); M. S. Chorghade (USA); P. W. Erhardt (USA); J. Fischer (Hungary); A. Ganesan (UK); G. Gaviraghi (Italy); T. Kobayashi (Japan); M. N. Liebman (USA); P. Lindberg (Sweden); Y. Martin (USA); P. Matyus (Hungary); A. Monge (Spain); T. J. Perun (USA); F. Sanz (Spain); J. Senn-Bilfinger (Germany); N. J. de Souza (India); G. Tarzia (Italy); H. Timmerman (Netherlands); M. Varasi (Italy); Z.-J. Yao (China).

ACKNOWLEDGMENTS

The authors are grateful to the following individuals for their valuable support, comments or suggestions: Koen Augustyns (University of Antwerp, Belgium), John Comer (Sirius, UK), Gunda Georg (University of Minnesota, USA), William J. Greenlee (MedChem Discovery Consulting, USA), Trevor Grinter (GSK, UK), Philip Jones (Consultant, UK), Hugo Kubinyi (Germany), John Macor (BMS, USA), Carlo Melchiorre (University of Bologna, Italy), J. Richard Morphy (Schering Plough Research, UK), Philip Portoghese (University of Minnesota, USA), Hans Ulrich Stiltz (Sanofi-Aventis, Germany), Hiromitsu Takayama (Chiba University, Japan), Antoni Torrens (Esteve Quimica, Spain), Shaomeng Wang (University of Michigan, USA), Camille G. Wermuth (Prestwick Chemical, France).

REFERENCES

1. C. G. Wermuth, C. R. Ganellin, P. Lindberg, L. A. Mitscher. *Pure Appl. Chem.* **70**, 1129 (1998).
2. IUPAC. *Compendium of Chemical Terminology*, 2nd ed. (the "Gold Book"). Compiled by A. D. McNaught and A. Wilkinson. Blackwell Scientific Publications, Oxford (1997). XML on-line corrected version: http://dx.doi.org/10.1351/goldbook (2006–) created by M. Nic, J. Jirat, B. Kosata; updates compiled by A. Jenkins.
3. J. Fischer, C. R. Ganellin (Eds.). *Analogue-based Drug Discovery*, Wiley-VCH, Weinheim (2006).
4. M. Dean, Y. Hamon, G. Chimini. *J. Lipid Res*. **42**, 1007 (2001).
5. G. Bringmann, A. J. P. Mortimer, P. A. Keller, M. J. Gresser, J. Garner, M. Breuning. *Angew. Chem., Int. Ed.* **44**, 5384 (2005).
6. M. W. Sinz. *Ann. Rep. Med. Chem*. **43**, 405 (2008).
7. R. Ng. *Drugs from Discovery to Approval*, Wiley-Blackwell (2009).
8. J. Kelder, P. D. J. Grootenhuis, D. M. Bayada, L. P. C. Delbressine, J. P. Ploemen. *Pharm. Res*. **16**, 1514 (1999).
9. S. A. Hitchcock, L. D. Pennington. *J. Med. Chem.* **49**, 7559 (2006).
10. M. Nordberg, J. H. Duffus, D. M. Templeton. *Pure Appl. Chem.* **79**, 1583 (2007).
11. T. P. Begley. *Nat. Chem. Biol*. **1**, 236 (2005).
12. Pubchem; http://pubchem.ncbi.nlm.nih.gov/
13. F. K. Brown. *Ann. Rep. Med. Chem*. **33**, 375 (1998).
14. H. Kubinyi. *Chemogenomics in Drug Discovery*, www.kubinyi.de/schering58-2006.pdf
15. K. G. Carson, B. D. Jaffee, G. C. B. Harriman. *Ann. Rep. Med. Chem*. **39**, 149 (2004).
16. G. M. Downs, P. Willett, W. Fisanick. *J. Chem. Inf. Comput. Sci.* **34**, 1094 (1994).
17. R. D. Brown, Y. C. Martin. *J. Chem. Inf. Comput. Sci.* **36**, 572 (1996).
18. D. Gorse, A. Rees, M. Kaczorek, R. Lahana. *Drug Discovery Today* **4**, 257 (1999).
19. W. M. Lau, A. W. White, S. J. Gallagher, M. Donaldson, G. McNaughton, C. M. Heard. *Curr. Pharm. Des.* **14**, 794 (2008).
20. F. W. Schueller. *Chemobiodynamics and Drug Design*, p. 405, McGraw-Hill, The Blakiston Division (1960).
21. C. G. Wermuth. *Drug Discovery Today* **11**, 348 (2006).
22. J. Owens. *Nat. Rev. Drug Discovery* **6**, 187 (2007).
23. A. C. Cheng, R. G. Coleman, K. T. Smyth, Q. Cao, P. Soulard, D. R. Caffrey, A. C. Salzberg, E. S. Huang. *Nat. Biotech.* **25**, 71 (2007).
24. T. L. Doan, M. Pollastri, M. A. Walters, G. I. Georg. *Ann. Rep. Med. Chem*. **46**, 385 (2011).
25. J. Proudfoot, O. Nosjean, J. Blanchard, J. Wang, D. Besson, D. Crankshaw, G. Gauglitz, R. Hertzberg, C. Homon, L. Llewellyn, R. Neubig, L. Walker, P. Villa. *Pure Appl. Chem*. **83**, 1129 (2011).
26. J. H. Duffus, M. Nordberg, D. M. Templeton *Pure Appl. Chem*. **79**, 1153 (2007).

27. K. J. Box, J. E. Comer. *Curr. Drug Metab*. **9**, 869 (2008).
28. K. J. Box, J. E. Comer, T. Gravestock, M. Stuart. *Chem. Biodivers.* **6**, 1767 (2009).
29. H. Gamsjäger, J. W. Lorimer, P. Scharlin, D. G. Shaw. *Pure Appl. Chem.* **80**, 233 (2008).
30. S. D. Pickett, C. Luttmann, V. Guerin, A. Laoui, E. James. *J. Chem. Inf. Comput. Sci.* **38**, 144 (1998).
31. M. J. McGregor, S. M. Muskal. *J. Chem. Inf. Comput. Sci.* **39**, 569 (1999).
32. H. Zhao, Z. Guo. *Drug Discovery Today* **14**, 516 (2009).
33. J. A. Di Massi, L. A. Faden. *Nat. Rev. Drug Discovery* **10**, 23 (2011).
34. M. Congreve, C. W. Murray, R. Carr, D. C. Rees. *Ann. Rep. Med. Chem*. **42**, 431 (2007).
35. J. B. Baell, G. A. Holloway. *J. Med. Chem*. **53**, 2719 (2010).
36. Medicines and Healthcare Products Regulatory Agency. *Good Laboratory Practice: Background and Structure*, http://www.mhra.gov.uk/Howweregulate/Medicines/Inspectionandstandards/GoodLaboratoryPractice/Structure/index.htm
37. Medicines and Healthcare Products Regulatory Agency. *Good Manufacturing Practice*, http://www.mhra.gov.uk/Howweregulate/Medicines/Inspectionandstandards/GoodManufacturingPractice/index.htm
38. IUPAC Green Chemistry Directory. http://www.incaweb.org/transit/iupacgcdir/overview.htm
39. R. F. Rekker. *Hydrophobic Fragmental Constant* (Pharmacochemistry library), Elsevier Scientific (1977).
40. The IUPAC International Chemical Identifier (InChI); http://IUPAC.org/inchi/
41. Solubility definitions; http://sirius-analytical.com; Application note 08/12
42. A. L. Hopkins, C. R. Groom, A. Alex. *Drug Discovery Today* **9**, 430 (2004).
43. P. D. Leeson, B. Springthorpe. *Nat. Rev. Drug Discovery* **6**, 881 (2007).
44. B. G. Katzung. In *Basic and Clinical Pharmacology*, 9th ed., Lang Medical Books, McGraw-Hill (2004).
45. K. A. Loomis, G. J. Brock. *Ann. Rep. Med. Chem*. **46**, 351 (2011).
46. H. van de Waterbeemd, B. Testa. *Adv. Drug Res.* **16**, 85 (1987).
47. H. van de Waterbeemd, R. E. Carter, G. Grassy, H. Kubiny, Y. C. Martin, M. S. Tute, P. Willett. *Pure Appl. Chem*. **69**, 1137 (1997).
48. Chronic Lymphocytic Leukemia Research Consortium Glossary. http://cll.ucsd.edu/glossary/glossary_m.html
49. T. T. Wager, Xinjun Hou, P. R. Verhoest, A. Villalobos. *ACS Chem. Neurosci.* **1**, 435 (2010).
50. G. W. Bemis, M. A. Murcko. *J. Med. Chem*. **39**, 2887 (1996).
51. Australian Academy of Science. *Good Prospects Ahead for Data Mining*, http://www.science.org.au/nova/050/050glo.htm
52. U.S. Food and Drug Administration home page. http://www.fda.gov
53. NIH Center for Macromolecular Modeling and Bioinformatics. *Nuclear Hormone Receptors*, http://www.ks.uiuc.edu/Research/pro_DNA/ster_horm_rec/
54. D. Maclean, J. J. Baldwin, V. T. Ivanov, Y. Kato, A. Shaw, P. Schneider, E. M. Gordon. *Pure Appl. Chem*. **71**, 2349 (1999).
55. S. C. Smith. In *Analogue-based Drug Discovery II*, J. Fischer, C. R. Ganellin (Eds.), p. 83, Wiley-VCH, Weinheim (2010).
56. W. J. Egan. *Ann. Rep. Med. Chem*. **42**, 449 (2007).
57. National Center for Biotechnology Information. *Just the Facts: A Basic Introduction to the Science Underlying NCBI Resources*. http://www.ncbi.nlm.nih.gov/About/primer/pharm.html
58. J. Kotz. *SciBX* **5** (15) (2012); http://dx.doi.org/10.1038/scibx.2012.380
59. J. Kelder, P. D. J. Grootenhuis, D. M. Bayada, L. P. C. Delbressine, J.-P. Ploemen. *Pharmaceutical Res.* **16**, 1514 (1999).
60. D. J. W. Grant. "Theory and origin of polymorphism", in *Polymorphism in Pharmaceutical Solids*, H. G. Brittain (Ed.), pp. 1–34, Marcel Dekker, New York (1999).

61. B. E. Evans, K. E. Rittle, M. G. Bock, R. M. DiPardo, R. M. Freidinger, W. L. Whitter, G. F. Lundell, D. F. Veber, P. S. Anderson, R. S. Chang. *J. Med. Chem.* **31**, 2235 (1988).
62. Worldwide Protein Databank. http://www.wwpdb.org
63. B. Fermini, A. F. Fossa. *Ann. Rep. Med. Chem*. **39**, 323 (2004).
64. C. Lipinski, F. Lombardo, B. W. Dominy, P. J. Feeney. *Adv. Drug. Deliv. Res*. **23**, 3 (1997).
65. D. C. Rees, M. Congreve, C. W. Murray, R. Carr. *Nat. Rev*. **3**, 660 (2004).
66. H.-J. Böhm, A. Flohr, M. Stahl. *Drug Discovery Today: Technologies* **1**, 217 (2004).
67. B.-M. Shamoon, C. Reinhard. *Ann. Rep. Med. Chem*. **38**, 261 (2003).
68. Daylight Chemical Information Systems home page. http://www.daylight.com
69. B. Gomes, D. de Graaf. *Ann. Rep. Med. Chem*. **42**, 393 (2007).
70. G. L. Patrick, Organic Chemistry *Instant Notes Series. Chemistry Series*, 2nd ed., Taylor & Francis Routledge, London (2004).
71. R. P. Sheridan, S. K. Kearsley. *J. Chem. Inf. Comput. Sci.* **35**, 310 (1995).

ANNEX: ACRONYMS USED IN MEDICINAL CHEMISTRY LITERATURE

ADMET	absorption, distribution, metabolism, excretion, toxicology
ABC	ATP binding cassette
BBB	blood–brain barrier
CRO	contract research organisation
DOS	diversity-oriented synthesis
EC	effective concentration
ED	effective dose
GLP	good laboratory practice
GMP	good manufacturing practice
GPCR	G-protein-coupled receptor
HTS	high-throughput screening
InChI	International Chemical Identifier
IND	investigational new drug
LE	ligand efficiency
LLE	ligand lipophilic efficiency
MDR	multidrug resistance
MPO	multiparameter optimisation
MTDL	multitarget-directed ligand
MTDD	multitarget drug discovery
PgP	P-glycoprotein
PDB	protein data bank
PPI	protein–protein interaction
PET	positron emission tomography
SAR	structure–activity relationship
siRNA	small inhibitory double-stranded RNA
SMILES	simplified molecular input line entry system
UHTS	ultra-high-throughput screening
V_d	volume of distribution

