FACULDADE DE MEDICINA
UNIVERSIDADE DO PORTO

MESTRADO INTEGRADO EM MEDICINA

2013/2014

César Augusto da Silva Portela

The Translational Approach between Computational Chemistry and Clinical Expertise in Drug Development

março, 2014

FACULDADE DE MEDICINA
UNIVERSIDADE DO PORTO

César Augusto da Silva Portela

The Translational Approach between Computational Chemistry and Clinical Expertise in Drug Development

Mestrado Integrado em Medicina

Área: Farmacologia e Terapêutica

Trabalho efetuado sob a Orientação de:
Doutor Patrício Soares da Silva

Trabalho organizado de acordo com as normas da revista:
Drug Discovery Today

março, 2014

U.PORTO
FMUP FACULDADE DE MEDICINA UNIVERSIDADE DO PORTO
UNIDADE CURRICULAR PROJETO DE OPÇÃO DISSERTAÇÃO / MONOGRAFIA

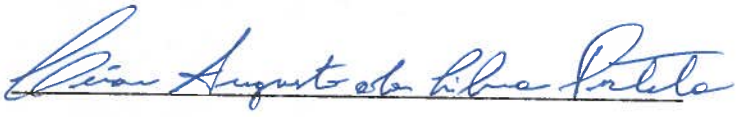
Projeto de Opção do 6º ano - DECLARAÇÃO DE INTEGRIDADE

Eu, César Augusto da Silva Portela, abaixo assinado, nº mecanográfico 199402741, estudante do 6º ano do Ciclo de Estudos Integrado em Medicina, na Faculdade de Medicina da Universidade do Porto, declaro ter atuado com absoluta integridade na elaboração deste projeto de opção.

Neste sentido, confirmo que **NÃO** incorri em plágio (ato pelo qual um indivíduo, mesmo por omissão, assume a autoria de um determinado trabalho intelectual, ou partes dele). Mais declaro que todas as frases que retirei de trabalhos anteriores pertencentes a outros autores, foram referenciadas, ou redigidas com novas palavras, tendo colocado, neste caso, a citação da fonte bibliográfica.

Faculdade de Medicina da Universidade do Porto, 07/03/2014

Assinatura conforme cartão de identificação:



U.PORTO
FMUP FACULDADE DE MEDICINA UNIVERSIDADE DO PORTO
UNIDADE CURRICULAR PROJETO DE OPÇÃO DISSERTAÇÃO / MONOGRAFIA

Projecto de Opção do 6º ano – DECLARAÇÃO DE REPRODUÇÃO

NOME

César Augusto da Silva Portela

CARTÃO DE CIDADÃO OU PASSAPORTE (se estrangeiro)	E-MAIL	TELEFONE OU TELEMÓVEL
10799186	cesarsportela@gmail.com	919063377

NÚMERO DE ESTUDANTE	DATA DE CONCLUSÃO
199402741	04/03/2014

DESIGNAÇÃO DA ÁREA DO PROJECTO

Farmacologia e Terapêutica

TÍTULO DISSERTAÇÃO/MONOGRAFIA (riscar o que não interessa)

The Translational Approach between Computational Chemistry and Clinical Expertise in Drug Development

ORIENTADOR

Professor Doutor Patrício Soares da Silva

COORIENTADOR (se aplicável)

É autorizada a reprodução integral desta Dissertação/Monografia (riscar o que não interessa) para efeitos de investigação e de divulgação pedagógica, em programas e projectos coordenados pela FMUP.

Faculdade de Medicina da Universidade do Porto, 07/03/2014

Assinatura conforme cartão de identificação: 

The Translational Approach between Computational Chemistry and Clinical Expertise in Drug Development

César Portela [1,2], Patrício Soares-da-Silva [3]

[1] Instituto Superior de Ciências da Saúde – Norte, Rua Central de Gandra, 1317, 4585-116 Gandra, Portugal

[2] REQUIMTE/Departamento de Química, Faculdade de Ciências, Universidade do Porto, Rua do Campo Alegre 687, 4169-007 Porto, Portugal

[3] Departamento de Farmacologia e Terapêutica, Faculdade de Medicina da Universidade do Porto, Alameda Prof. Hernâni Monteiro, 4200- 319 Porto, Portugal

Corresponding author

Portela, C. (cesarsportela@gmail.com)

Keywords

Drug development, Computational chemistry, Clinical expertise, Translational research

Teaser phrase

How the clinical practice and expertise may associate with the computational drug development methods in a more effective definition of new molecules with potential therapeutic effect.

Research highlights

Drug development by pharmaceutical companies generally depends of commercial interests and market opportunities, limiting the investment in innovation.

Academic research lacks funding and acts on specific fields, without receiving input from clinicians or the industry.

Clinical needs and therapeutic problems are not always a priority in drug development.

Clinical expertise is not taken into account in project creation and definition of the future therapeutic viability of newly designed drugs.

Translational research by combining clinical practice, applied research and computational chemistry can surpass limitations at present.

Abstract

Traditionally, the first step in the development of new drugs is the definition of the target, by choice of a biological structure involved in a disease or by the recognition of a molecule with some degree of a biological activity that presents itself as druggable and endowed with therapeutic potential. The complexity of the pathophysiological mechanisms of disease and of the structures of the molecules involved creates several challenges in this drug discovery process. These difficulties also come from independent operation of the different parts involved in drug development, with little interaction between clinical practitioners, academic institutions and large pharmaceutical companies. Generally, research in this area is purpose specific, performed by specialized researchers in each field, without major inputs from clinical practitioners on the relevance of such strategy for future therapies. Translational research is a path of shifting the way these relationships operate towards a process in which new therapies can be generated by linking experimental discoveries directly to unmet clinical needs. Computational chemistry methods provide valuable insights on experimental findings and pharmacological and pathophysiological mechanisms, allow the virtual construction of new possibilities for the synthesis of new molecular entities, and pave the way for informed cost-effective decisions on expensive research projects. This review focus on the current computational methods used in drug design, how they can be used in a translational research model that starts from clinical practice and research-based theorization by medical practitioners and moves to applied research in a computational chemistry setting, aiming the development of new drugs for clinical use.

Introduction

The majority of drugs available today come from different approaches, having in common the following development steps: target identification and validation, lead identification, lead optimization and non-clinical trials. The complete process generates active molecules that are evaluated in clinical trials before being subjected to approval as new drugs for disease treatment [1]. Traditionally, the method consists in constructing hypotheses on a molecular component of a particular mechanism in a specific pathology, theorizing on how to overcome a disease-based pathophysiological mechanisms and finding small molecules to deliver the corrective solution. The method also uses another common aspect of these approaches that has been the traditional concept of a “receptor” as a target [2].

Strategies in discovering small molecules that can fulfill the hypothesis have varied over the years, with isolation of lead compound from plants and animals, use of empirical chemistry and applied pharmacology, development of rational drug design based on new knowledge in physiology and pathophysiology and drug repositioning [2]. The isolation of lead compounds from plants and animals delivered some of the most potent and widely used drugs today. Mankind has been using natural products for therapeutic purposes for a very long time. The extracts from plants and animals usually contain a mixture of ingredients, either beneficial or adverse. Early drug development focused on identifying the entities responsible for the beneficial effects and purifying them. The use of a single molecule facilitates the evaluation of safety and efficacy, contrary to the use of a complex mixture. Drug development based on extracts from animals and plants is performed mainly by two methods: the research of ethnic remedies looking for evidence of a therapeutic effect and the screening of different extracts of

plant and animal parts against batteries of biological and genomic test systems looking for a potentially interesting biological action. This strategy creates several difficulties as adequate drug quantities obtained by chemical extraction are often a limiting factor. Also, the search of new molecules is frequently developed without a strategy involving prior clinical needs, the definition of underlying pathophysiological mechanisms to which a disease can be treated or the delineation of an eventual future place in therapy. Additionally, the entities obtained from plants or animals are often large and complex, making them difficult to synthesize [3]. The use of empirical chemistry coupled with applied pharmacology is one of the most productive sources in drug development. The identification of a pharmacophore consists in the design of a simple molecule with similar pharmacological activity. The synthesized entity can then serve as a model for further modifications to improve pharmacokinetics and pharmacodynamics. This process is lengthy, strenuous and expensive, involving synthesis of a range of related compounds, molecule purification, and structure characterization, pharmacological and toxicological properties testing [2]. The advent of combinatorial chemistry and high-throughput screening (HTS) allowed that a huge number of related molecules could be produced in lesser time. This approach depends on automation to synthesize and screen a high number of molecules to find all those that can enable a desired biological action. This strategy has the advantage of requiring minimal compound design or minimal prior pathophysiological and pharmacological knowledge. Although the technologies required in screening large libraries of compounds have become more efficient, the development of suitable systems in which compounds are tested is still challenging and the methods are expensive. Furthermore, although traditional HTS often results in multiple hit compounds, some of which are capable of being modified into a lead and

later a novel therapeutic, the hit rate for HTS is often extremely low [4]. Again, the development is mostly made without taking into account the clinical needs [5].

The use of rational drug design based on new knowledge in physiology and pathophysiology is one of the main areas in which the clinical practitioner can take a role in drug development. Our understanding of physiology and pathophysiology has improved substantially and there has been an increase in accuracy of technologies available for drug design. Clinical practitioners should be able to develop and evaluate a novel research proposal aiming the characterization of disease mechanisms and determine its potential applicability and value as a therapeutic intervention towards different putative targets. These putative targets should then be evaluated *in silico* on their properties. The use of computational chemistry allows the prediction of the structure of the binding site of a receptor in three dimensions from its amino-acid sequence. Based on this information, it is possible to virtually design groups of molecules that may bind with high affinity to that site [6].

The drug repositioning strategy is the second of the main areas in which the clinical practitioner can be an active part in drug development. Several effective and lucrative drugs were repurposed, without their development being determined for their present indications. Serendipitous discovery of new pharmacological effects and their therapeutic applicability by different old drugs of the same therapeutic group have led to the implementation of new uses [7]. The identification of previously unknown pharmacological effects of a known drug and the identification of the nature of adverse effects of drugs in clinical practice can serve as basis for drug discovery. There is strong evidence that such off-target interactions, or polypharmacology, are common among many approved drugs [8]. The use of computational chemistry can be explanatory on how a single molecule can act on multiple targets, by definition of its form, size,

analogy with endogenous ligands or other drugs, charge distribution and complementarity with receptors. The chosen molecule can constitute a lead compound for further research. The fact that the first contact with these adverse effects comes from the physician may determine its choice where to pursue and where not to pursue, in view of its relevance for future therapy.

Although all of these strategies are currently in use by research teams in academic groups, biotechnology companies and pharmaceutical industries, they operate independently, each with its own objectives and methods. Furthermore, the problem of not supplying the needs present in the everyday clinical practice is as relevant as ever. It must be accepted by all entities involved that the traditional ways of developing drugs are becoming ineffective and cannot accompany the rapid developments in health sciences [9]. Studies in the field of drug discovery and development show that large pharmaceutical companies do not present themselves as leading examples in innovation, but have commercial interest as their major concern [10]. This also has repercussion in the development of new drugs based on the fraction of the market that a determined therapeutic group may achieve and not on the clinical needs. The public sector represented by academics and the biotech companies are becoming the main contributors in drug discovery [11]. The problem with these sectors is that academic researchers or small biotech companies are often not well-trained in clinical research. There is also the issue of lack of training in business strategies resulting in little access to the necessary funding for generating research data attractive to investment.

The final point is that the lack of communication between all the referred parties has resulted in many valid ideas not being developed in research and many drug researches being unproductive. A new model for the development of new drugs is emerging called translational research and represents a more focused strategy for

creating new drugs than the traditional model [12]. The basic concept consists in combining the needs of patients with research-originated concepts provided by clinical practitioners and with state-of-the-art data on the subject as the basis for planning new therapies.

In this review we will discuss how translational research may be applied in combining patient and clinical unmet needs with computational drug discovery based on clinical expertise.

Translational research and the use of computational strategies

The definition of translational research is not consensual, with multiple definitions for its meaning and its use [13]. The expression translational research provided here is based on the notion that the development of new drugs must relate directly to patient needs and that could be performed by coupling computational and laboratory research with observations originated in clinical practice. The achievement of the translational approach in drug design is the incorporation of a specific clinical need from the beginning of the research process. The traditional research-based drug design is based on applying data from basic cellular mechanisms to the development of new therapies. Translational research encompasses this concept, with the advantage of targeting mechanisms underlying clinically relevant problems and developing molecules with potential action over those issues directly. Translational research covers the main components that should be involved in drug development: clinical practice and expertise, laboratory investigation, and health benefits in society [14]. The process involves two main stages, being one of them the connection between clinicians and applied research, and named T1.The other stage is named T2 and corresponds to the

connection between clinicians and community [13]. The T1 concept is mainly performed in universities or other institutes of higher education, and focus on the laboratory discoveries that relate to specific clinical endpoints. The proximity between clinical departments of a central hospital and the academic researchers of the associated faculty or university enables laboratory scientists and practicing physicians to gather and provide the discussion on how clinical practices and laboratory data can be applied in drug design for different diseases. The unmet needs among patients and the quantitative and qualitative different response to existent drugs, recorded by physicians, can be shared with laboratory researchers. This communication allows the planning of potential solutions and the creation of new projects based on the prior knowledge of the underlying molecular mechanisms of diseases and drugs. The T2 concept integrates community outreach programs with clinical practices, with the aim of providing a means for understanding how well treatment strategies are working at a population level. This notion may also allow the identification of needs of patients and the quantitative and qualitative different response to existent drugs for posterior debate and consequent research [14].

The effective communication and regular collaboration between all the involved parts are the basis of translational research and can facilitate the interaction between clinicians that treat patients and computational chemistry scientists that could explore the data provided by the first. The clinicians would provide patient and clinical issues in need of solution, input in diseases lacking therapeutic options, applicable pathophysiological mechanisms for drug design directed to treatment of yet unsolved problems, interesting drug actions and adverse effects or patient responses to treatments. The information thus provided would serve as a starting point in the definition of biological targets and lead compounds, being then applied to the virtual design of

molecules for further selection based on activity prediction, by use of different computational tools. The predicted potentially active molecules could then be synthesized and biologically tested (Figure 1).

Computational methods are capable of increasing the rate of discovery of hit compounds because it uses a much more targeted approach. It has the advantages of attempting to explain the molecular basis of a therapeutic activity and the prediction of possible derivatives that could improve activity [4]. There are many computational strategies applicable to drug design. One way to classify these methods is by categorizing them as either “Ligand-based methods”, where discovery of opportunities initiates from knowledge about small molecules and their action, or “Structure-based methods”, where discovery initiates from knowledge about macromolecules involved in a disease pathophysiology or symptomatology. The approaches in drug design using ligand-based methods can be systematized in 4 main categories: activity and chemical similarity, adverse effects similarity, indication reallocation and shared molecular pathology. As for the approaches in structure-based methods, its main basis consists in pathophysiological mechanism definition, although the concept of shared molecular pathology can also be applied. The referred systematization can be described as followed:

- Activity and chemical similarity: The structure and chemical properties of a molecule correlate with its pharmacological action. The study of shared physicochemical characteristics between molecules presenting the same biological activity allows the definition of criteria for new molecule design. This concept is named quantitative structure-activity relationships (QSAR) and constitutes a rational basis for drug development [15]. This concept is still quite valid and useful in designing molecules based on an endogenous ligand of

known structure, to develop agonists or antagonists of its activity. The same approach can be applied to previously approved drugs, designing new ones to overcome pharmacokinetic problems or improve efficacy and safety. It can also be quite helpful when combined with the concept of adverse effects similarity, in the explanation and improvement of an interesting action, for the design of new drugs.

- Adverse effect similarity: Existent drugs can be correlated to clinical effects through their adverse effects, which represent unintended biological actions of the active molecule. The unintended actions can be beneficial in a determined disease condition, posing a possibility of a treatment that can be further researched. Adverse effects also provide a means to connect drugs between themselves to establish QSAR or a pharmacophore, even in cases where the precise pharmacological mechanism of the adverse effect is unknown. Adverse effects also provide a means to connect drugs to diseases. The manifestation of an adverse effect can be similar to that of a disease, raising the possibility that the underlying physiological process may be similarly disturbed by both the drug and the disease pathophysiological mechanism [16].
- Indication reallocation: The knowledge of drug indications for disease is a tool for definition of new lead compounds. Diseases can be considered similar if they share a significant number of drugs in the established therapeutic regimens. In each pair of similar diseases, the drugs that are currently used against only one of the diseases can then be considered as candidates as drugs for the other disease in the pair [17]. One of these drugs can be defined as a lead compound. The lead compounds found can serve as a basic structure for the design of new structures that share chemical similarities.

- Shared molecular pathology: the existence of some common aspect of underlying molecular pathophysiology between two diseases allows that a drug which presents a known pharmacological mechanism can be repositioned from one indication to another. This strategy allows drug repurposing and also the definition of new lead compounds [18].
- Pathophysiological mechanisms: the definition of the macromolecules involved in a disease and the way each one is affected is one the basis of the rational drug design. The definition of the tridimensional structure of the selected macromolecules, by x-ray crystallography or other experimental method, allows their use in the design of virtual molecules and to predict their ability to associate with the active site that may result in a biological action with potential usefulness in therapy [19].

Ligand-based approaches might be preferred, if there is interest to understand more precise pharmacological properties, or if rich pharmacological and chemical data for drugs or endogenous small molecules is available. Structure-based approaches may be preferred when the purpose is to focus on a specific disease. While each of these approaches present unique informatic challenges, successful strategies often incorporate elements from both methods [20]. There are several programs created for the purposes and techniques referred for ligand-based and structure-based methods of drug design. The number of computational tools applicable in drug discovery campaigns suggests that there are no fundamentally superior techniques, but the performance of methods varies greatly with target protein, available data, and available resources [4]. Although effective in their function, there are situations where there is the need of a prior or further study of the receptors selected, of the lead compounds and of the new molecules designed by one or both methods. It is common to use software for molecular

mechanics and dynamics simulations, quantum mechanics calculations, absorption-distribution-metabolism-excretion/toxicity (ADME/T) predictions, molecular visualization and chemoinformatics, each with its own applicable features in drug design (Box 1).

Ligand-based methods for drug design

Ligand-based methods are based on the principle which states that similar chemical structures tend to present similar biological activities [21]. These methods rely on prior knowledge of biological ligands or prior drugs and macromolecular structures, generally not being applicable in cases where no ligands for a given putative receptor exist. The main methods are 3D pharmacophore modeling and QSAR.

3D pharmacophore modeling can be used in the absence of a receptor structure. The prerequisite is the condition of having a set of known ligands representative of essential ligand–macromolecule interactions from which can be extracted the common chemical features from their 3D structures. IUPAC defines pharmacophore as “an ensemble of steric and electronic features that is necessary to ensure the optimal supramolecular interactions with a specific biological target and to trigger (or block) its biological response” [22]. The common chemical characteristics that are usually selected are the presence of hydrogen-bond acceptors, hydrogen-bond donors, hydrophobic regions and positively or negatively charged groups. A 3D pharmacophore can also be derived from a receptor structure by observing the interactions between macromolecule and ligand. As such, shape and excluded volume information can be added to the pharmacophore. This has the advantage of designing molecules that not only have the selected binding features but can also predictively fit into the active site.

By definition, a pharmacophore is based on the concept of similarity between ligands, with the definition of an essential backbone for activity and all the substituents that can determine the physicochemical properties that lead to biological activity. The concept of pharmacophore has found widespread use in hit-and-lead identification and also in following lead optimization, being very successful in drug discovery [23]. 3D pharmacophore generation from a set of ligands involves two main steps. The first one corresponds to the definition of the conformations of each ligand most probably involved in the interaction with the receptor. The second one is the alignment of the multiple ligands (in their selected conformations) to determine the common chemical features needed to design a 3D pharmacophore. There are two types of pharmacophore models. One is the 3D model based on QSAR that is established with a relationship with the degrees of activity. The most common model involves a training set with only active ligands. The new compounds can be estimated qualitatively by whether they match the established 3D model [24]. The application of the 3D pharmacophore technique is demonstrated by the work in which were developed CB1 cannabinoid receptor antagonists for obesity treatment. Unified pharmacophore models for the CB1 receptor ligands were developed by incorporation into the superimposition model for the known cannabinoid agonists. From this information it was possible to design antagonists by introducing aromatic rings for steric hindrance [25]. The success of this approach came from the application of the concept of activity and molecular similarity, using a 3D pharmacophore definition method.

QSAR modeling is also an established method, being used as a computational tool for rationalizing and correlating physicochemical properties with experimental binding data or inhibitory activity of chemical compounds [26]. QSAR consists mainly in two different techniques, 2D and 3D QSAR. 2D QSAR consists in defining an

equation that can be used to predict activity based on descripted physicochemical properties of a compound. The equation is a correlation between a set of independent variables (chemical descriptors) and a dependent variable such as receptor binding ability for the compound of interest. The equation is established and applied using algorithms like regression-analysis algorithms, multivariate analysis algorithms, heuristic algorithms or genetic algorithms [27]. 3D QSAR is a QSAR approach based on a set of predefined 3D molecular structures. The molecular descriptors used contain physicochemical properties and conformational coordinate-derived information. This technique uses a 3D grid of points around the molecule, each point having properties associated with it that can vary in a field-like manner from point to point, such as steric interactions or electrostatic potential. Therefore, this method can be used for predicting the binding capability of a ligand to the active site of a specific receptor. The construction of the 3D-QSAR model needs a training set, containing at least 20 active compounds with activity over the selected pathophysiological mechanism. The next step is to generate conformations and alignments of the training set molecules. A dimensionality reduction step is then inserted to extract the features of the 3D interaction field that are most strongly determining the activity before the actual predictive model is built. At last, a test set with some known active molecules is used to examine the prediction ability of the built 3D QSAR model [28]. The applicability of the QSAR method is exemplified in a research work that involved the computational design approach to screen biomaterials with anti-atherogenic efficacy. Several amphiphilic macromolecules were quantified in terms of 2D and 3D descriptors. QSAR models with the referred descriptors for anti-atherogenic activity were constructed by screening a total of 1164 parameters against the corresponding, experimentally measured potency of inhibition of oxidized LDL uptake in human monocyte-derived

macrophages. Five key descriptors were identified to provide a strong linear correlation between the predicted and observed anti-atherogenic activity values, and were then used to correctly forecast the efficacy of three newly designed biomaterials. Thus, a new ligand-based drug design framework was successfully adapted to computationally screen and design biomaterials with cardiovascular therapeutic properties [29]. The research presented is a good example of translational research, involving a clinical need in atherogenesis prevention, computational chemistry and biomaterials research.

Ligand-based methods can be used to determine minimal and common structures predictively responsible for biological activity. The data needed to start a research program is the identification of endogenous ligands or exogenous compounds that present the same activity (activity and chemical similarity, adverse effects similarity, indication reallocation). This information can be provided by clinicians of a designed specialty, bearing in mind the therapeutic relevance of the data. The 3D structure of the selected molecules can be used to establish a 3D pharmacophore or QSAR models for further design of new compounds with potential pharmacological action. Several other examples could be presented, with different strategies of approach, from activity and molecular similarity, adverse effects similarity, indication reallocation and shared molecular pathology, as seen before. One of these examples is the adverse effects similarity presented by cyclobenzaprine, which reportedly originated serotoninergic syndrome. A virtual screening provided evidence that cyclobenzaprine blocks, with moderate to high potency, the serotonin and norepinephrine transporters as well as five serotonin receptor subtypes at therapeutically relevant concentrations [30].

Structure-based methods for drug design

Structure-based methods are strategies that explore macromolecular structural information, combined with scoring functions, in order to predict ligand–receptor affinity. Ligands are defined as interaction partners for a given receptor. This concept has been recently reverted to dock one small molecule against a panel of multiple receptors [31].

Molecular docking is the preferred method to investigate how a ligand interacts with the receptor, when the structure of the target macromolecule is known. Molecular docking consists in an algorithm that determines how a molecule may establish connections in the binding site of a putative receptor and tries to predict the strength of the interaction. This method is an attempt of mimicry of the process of formation of a non-covalent complex by bringing together a macromolecular receptor and a ligand. The virtual complex obtained reveals the electrostatic and steric complementarity between the macromolecule and its different ligands. A docking algorithm performs an attempt of prediction of the correct positions of ligands at the binding site of a macromolecule and establishes a ranking of the obtained poses. The accomplishment of position prevision and accurate ranking is challenging, and so far none of the known docking programs were able to solve both of them perfectly. Prediction of possible binding positions in an active site is more straightforward, being performed by most programs. Because of its success at binding position prediction, docking is a well-established drug-design technology employed in structure-based methods [32]. The prerequisite for docking techniques and structure-based drug design is the existence of a 3D structure of a target, preferably in complex with a ligand. The 3D structure may be a crystallographic x-ray structure or an NMR structure. The structure of the ligand or of known active drugs can lead to the design of new molecules, based on the previously described ligand-based methods. The observation of the form, size, charge and

electrostatic potential distribution of the active site can also lead to the design of new virtual compounds. Once an appropriate set of molecular candidates has been designed, they can be docked into the active site allowing a further reduction of the number of hits based on the scoring functions. The docking results are examined visually or submitted to further computational calculations to choose candidates for synthesis and biological assays [33]. An example of the referred sequence of research work is the design of a series of coumarins to act as TNF-α converting enzyme inhibitors. The compounds were designed to bind in a pocket of the enzyme based on the docking study. Twelve analogues were synthesized and most of compounds were active in vitro, showing TNF-α converting enzyme inhibition as well as cellular TNF-α inhibition [34]. The prior definition of a pathophysiological mechanism allowed the definition of a target for drug development. The clinical importance of this intervention is demonstrated by the fact that overproduction of TNF-α is responsible for many autoimmune disorders such as rheumatoid arthritis, psoriasis, Crohn's disease, ulcerative colitis, among others. The clinical success of anti-TNF-α biologic agents for treating inflammatory diseases, such as infliximab or adalimumab, have confirmed that inhibition of TNF-α is an important approach for an effective treatment for several autoimmune diseases [35, 36]. Their use permitted overcoming a clinical need and an important health problem in populations, as is intended in translational research.

Homology modeling is a useful approach to develop structure-based drug design when the 3D model of a target protein is needed and whose structural configuration is not experimentally determined. The requisites here are the availability of the sequence of its amino acids and the experimental determined 3D structure for one or more sufficiently proteins similar to the selected target. Homology modeling performs the assembling of a model of the target protein from its amino acid sequence using the

experimental 3D structures of related homologous proteins as templates [37]. The concept is based on the experience that similar amino-acid sequences lead to similar 3D topographies. The conservation of regions between the active site of the studied protein and the template structures gives good accordance [38]. The quality of a homology model is consequent to the quality of the chosen template structure and the sequence alignment performed, and is biased by low sequence identity between the target and the template. Models with more than 50% sequence identity are believed to be accurate enough for drug design application. In this range, the root-mean-square deviation between the experimental structure and the model may be around 1 Å, which is equivalent to the typical resolution of structures solved by NMR. In the 25–50% identity range, errors can be more severe and are frequently located in the flexible loops. The homology model can be used for the assessment of druggability and mutagenesis experiments, but should be applied with caution for drug design. Below 20–25% sequence identity, a model is usually not usable for drug design because serious errors can occur [37]. Homology modeling was used in the prediction of the 3D structure of the protein Rv3802c. Rv3802c is an essential cell wall lipase of Mycobacterium tuberculosis. The modeling of its structure for the first time provided insight in identifying the ligand binding sites and potential inhibitors effective towards mycobacterial proteins. Two diverse molecules have been identified as potential inhibitors effective towards Rv3802c by docking on the modelled macromolecule [39].

Structure-based methods can be used to study putative receptors involved in a pathophysiological process associated with a disease that constitutes a relevant problem in society (shared molecular pathology, pathophysiological mechanism definition). The importance of the disease can be determined by the team of physicians involved. The choice of the target in the pathophysiological process can also be determined by the

clinicians, supported by experimental evidence and clinical expertise. The selected macromolecules can be studied using structure-based methods to determine their conformation and configuration. The interaction with the proper endogenous ligand and with new potential drugs can also be simulated. This work allows the prediction of activity and the selection of the candidates for synthesis and activity evaluation. The discovery that raltegravir acts as a metnase inhibitor is an example on how structure-based methods can be used in drug repurposing and development. Metnase is a DNA repair enzyme which can constitute a potential target for adjuvant cancer therapy. Raltegravir was identified as a metnase inhibitor via structure-based virtual screening studies, being in fact confirmed that it presents the predicted action, at doses that are roughly ten times higher than the currently approved maximum dose [40].

Databases containing bioactivity records

The development in recent years of databases integrating diverse types of data such as structural data and drug adverse effects brought a powerful tool to drug design. The information that these databases carry was previously hardly accessible in electronic form at the public domain [41]. The databases differ in functionality, but have a common purpose of integrating different types of data. These databases may be just molecular structure collections, or provide relevant type of data, such as quantitative bioactivity of the molecules and their macromolecular targets, as well as data on targeted illnesses. Some of the available databases attempt to link small-molecule data, biological targets data and available assay data [42, 43]. There are millions of bioactivity data points available, which can be used for ligand-based or structure-based methods. The presentation of a clinical need in therapeutics can lead to a

search in these databases of compounds that show activity in a given problem. The selected compounds can be further studied by ligand-based methods to determine the minimal and common structure predictively responsible for activity, constituting the base for new drug design (activity and chemical similarity, adverse effects similarity, indication reallocation). In the same manner, the databases can provide information on putative receptors involved in a pathophysiological process and the definition of a common mechanistic ground with the subject presented by a team of physicians. The putative receptors can be studied using structure-based methods to determine their conformation and configuration, the process in which the interaction with the selected small molecules proceeds, and the simulation of interaction with new virtual molecules that could be developed to potentially active compounds (shared molecular pathology, pathophysiological mechanism definition). The presentation of a newly found adverse effect of a drug can start a selection of molecules that present the same action. The application of ligand-based methods allows the definition of the chemical properties of the different molecules presenting the same activity. New molecules can be designed presenting the molecular features determined as essential for activity (adverse effects similarity). The ZINC chemical library [44] is an example of a library used in a ligand-based similarity search, for the identification of potential anticancer compounds. The search was directed to the urokinase receptor. This receptor serves as a docking site to the serine protease urokinase-type plasminogen activator to promote extracellular matrix degradation and tumor invasion and metastasis. The search for inhibitors gave 127 derivatives that share the core structure of the molecules that act on the urokinase receptor. These derivatives were purchased and tested for inhibition of urokinase receptor binding to serine protease urokinase-type plasminogen activator. Cellular studies showed that compounds blocked invasion, migration and adhesion [45].

Future prospects

The combination of clinical practice and expertise with computational chemistry can be accomplished in the form of a translational discovery center. This center consists in an entity with a structure of research based on the creation of teams of clinicians of a given specialty, computational chemistry scientists and medicinal chemists, with the purpose of defining projects for drug development oriented to meet patient and clinical needs. The joining of knowledge can allow a more rational drug design, with a great input from clinicians in target definition and validation or lead compound selection. That creates the basis of research work from which new drug designs are pursued. The design obtained can be further developed in partnerships with different contributors. The concept may lead to a new style in the field of drug design. It has the potential to benefit all parties, pursuing the purpose of new and better drugs based on community needs and not simply on commercial interests. It can also provide academic researchers with access to funding and expertise from biotech and pharmaceutical companies, while providing opportunities for the pharmaceutical companies to access innovative research. This model for integrative drug development allows the potential funding and further development of research by connecting academia, industry, venture capital firms, philanthropic organizations, advocacy groups, independent consultants and contract research organizations. The concept is of a technology incubator for the design and possible creation of new effective drugs based on society concerns and clinical needs.

The success of this concept in drug discovery will depend on the effectiveness of communication of the parts involved and the willingness to prioritize research directed to aspects of disease and therapy that benefit the patient.

Translational research is a central new strategy in the field of drug development. The combination of clinical expertise and computational chemistry could be an effective way of applying the concept.

References

[1] Lindsay, M.A. (2003) Target discovery. *Nat. Rev. Drug Discov*. 2, 831–838

[2] David, G.T. (2011) Scientific process, pharmacology and drug discovery. *Curr. Opin. Pharm.* 11, 528–533

[3] Raza, M. (2006) A role for physicians in ethnopharmacology and drug discovery. *J. Ethnopharm.* 104, 297–301

[4] Sliwoski, G. et al. (2014) Computational Methods in Drug Discovery. *Pharmacol. Rev.* 66, 334–395

[5] Ratti, E. and Trist, D. (2001) Continuing evolution of the drug discovery process in the pharmaceutical industry. *Pure Appl. Chem.* 73, 67-75

[6] Kalyaanamoorthy, S. and Chen, Y.P. (2011) Structure-based drug design to augment hit discovery. *Drug Discov. Today* 16, 831-839

[7] Ashburn, T.T. and Thor, K.B. (2004) Drug repositioning: identifying and developing new uses for existing drugs. *Nat. Rev. Drug Discov.* 3, 673–83

[8] Keiser, M.J. et al. (2009) Predicting new molecular targets for known drugs. *Nature* 462, 175–181

[9] Cressey, D. (2011) Traditional drug-discovery model ripe for reform. *Nature* 471, 17–18

[10] Bennani, Y.L. (2011) Drug discovery in the next decade: innovation needed ASAP. *Drug Discov. Today* 16, 779–792

[11] Stevens, A.J. et al. (2011) The role of public-sector research in the discovery of drugs and vaccines. *N. Engl. J. Med*. 364, 535-541

[12] Fishburn, C.S. (2013) Translational research: the changing landscape of drug discovery. *Drug Discov. Today* 18, 487-494

[13] Woolf, S.H. (2008) The meaning of translational research and why it matters. *JAMA.* 299, 211–213

[14] Ledford, H. (2008) Translational research: the full cycle. *Nature* 453, 843–845

[15] Eckert, H. and Bajorath, J. (2007) Molecular similarity analysis in virtual screening: foundations, limitations and novel approaches. *Drug Discov. Today* 12, 225-233

[16] Campillos, M. et al. (2008) Drug target identification using side-effect similarity. *Science* 321, 263–266

[17] Chiang, A.P. and Butte, A.J. (2009) Systematic evaluation of drug-disease relationships to identify leads for novel drug uses. *Clin. Pharmacol. Ther.* 86, 507–510

[18] Suthram, S. et al. (2010) Network-based elucidation of human disease similarities reveals common functional modules enriched for pluripotent drug targets. *PLoS Comput. Biol.* 6, DOI:10.1371/journal.pcbi.1000662 (http://www.ploscompbiol.org/)

[19] Pérot, S. et al. (2010) Druggable pockets and binding site centric chemical space: a paradigm shift in drug discovery. *Drug Discov. Today* 15, 656-667

[20] Koutsoukas, A. et al. (2011) From *in silico* target prediction to multi-target drug design: Current databases, methods and applications, *J. Proteomics*, 74, 2554–2574

[21] Bender, A. and Glen R.C. (2004) Molecular similarity: a key technique in molecular informatics. *Org. Biomol. Chem.*2, 3204–3211

[22] Wermuth, G. et al. (1998) Glossary of terms used in medicinal chemistry (IUPAC Recommendations 1998). *Pure Appl. Chem.* 70, 1129–1143

[23] Leach, A.R. et al. (2010) Three-dimensional pharmacophore methods in drug discovery. *J. Med. Chem.* 53, 539–558

[24] Wolber, G. et al. (2008) Molecule-pharmacophore superpositioning and pattern matching in computational drug design. *Drug Discov. Today* 13, 23–29

[25] Shim, J. et al (2002) Molecular Interaction of the Antagonist N-(Piperidin-1-yl)-5-(4-chlorophenyl)-1-(2,4-dichlorophenyl)-4-methyl-1H-pyrazole-3-carboxamide with the CB1 Cannabinoid Receptor. *J. Med. Chem.* 45, 1447-1459

[26] Sprous, D.G. et al. (2010) QSAR in the pharmaceutical research setting: QSAR models for broad, large problems. *Curr. Top. Med. Chem.* 10, 619–637

[27] Tropsha, A. and Golbraikh, A. (2007) Predictive QSAR modeling workflow, model applicability domains, and virtual screening. *Curr. Pharm. Des*. 13, 3494–3504

[28] Clark, R.D. (2009) Prospective ligand- and target-based 3D QSAR: state of the art 2008. *Curr. Top. Med. Chem*. 9, 791–810

[29] Daniel, R. et al (2013) In silico design of anti-atherogenic biomaterials. *Biomaterials* 34, 7950-7959

[30] Gillman, P.K. (2009) Is there sufficient evidence to suggest cyclobenzaprine might be implicated in causing serotonin toxicity? *Am. J. Emerg. Med.* 27, 509–510

[31] Rognan, D. (2010) Structure-based approaches to target fishing and ligand profiling. *Mol. Inf.* 29, 176–187

[32] Brooijmans, N. and Kuntz, I.D. (2003) Molecular recognition and docking algorithms. *Annu. Rev. Biophys. Biomol. Struct.* 32, 335–373

[33] Alonso, H. et al. (2006) Combining docking and molecular dynamic simulations in drug design. *Med. Res. Rev.* 26, 531–568

[34] Yang, J.S et al. (2010) Structure based optimization of chromen-based TNF-a converting enzyme (TACE) inhibitors on S10 pocket and their quantitative structure–activity relationship (QSAR) study. *Bioorganic & Medicinal Chemistry* 18, 8618–8629

[35] Lipsky, P. E. et al. (2000) Infliximab and methotrexate in the treatment of rheumatoid arthritis. Anti-Tumor Necrosis Factor Trial in Rheumatoid Arthritis with Concomitant Therapy Study Group, *N. Eng. J. Med.* 343, 1594-1602

[36] Machold, K. P. and Smolen, J. S. (2003) Adalimumab - a new TNF-alpha antibody for treatment of inflammatory joint disease. *Expert Opin. Biol. Ther.* 3, 351-360

[37] Cavasotto, C.N. and Phatak, S.S. (2009) Homology modeling in drug discovery: current trends and applications. *Drug Discov. Today* 14, 676–683

[38] Baker, D. and Sali, A. (2001) Protein structure prediction and structural genomics. *Science* 294, 93–96

[39] Saravanan, P. et al. (2012) Targeting essential cell wall lipase Rv3802c for potential therapeutics against Tuberculosis, *J. Mol. Graphics Model.* 38, 235–242

[40] Oprea, T.I. and Matter, H. (2004) Integrating virtual screening in lead discovery. Curr. Opin. Chem. Biol. 8, 349–358

[41] Kuhn, M. et al. (2010) A side effect resource to capture phenotypic effects of drugs. *Mol. Syst. Biol.* 6, 343, DOI: 10.1038/msb.2009.98 (http://www.nature.com/msb)

[42] Liu, T. et al. (2007) BindingDB: a web-accessible database of experimentally determined protein–ligand binding affinities. *Nucleic Acids Res.* 35, DOI: 10.1093/nar/gkl999 (http://nar.oxfordjournals.org/)

[43] Hendlich, M. et al. (2003) Relibase: design and development of a database for comprehensive analysis of protein–ligand interactions. *J. Mol. Biol.* 326, 607–620

[44] Irwin, J. J. and Shoichet, B. K. (2005) ZINC − A Free Database of Commercially Available Compounds for Virtual Screening. *J. Chem. Inf. Model.* 45, 177-182

[45] Wang, F. et al. (2012) Design, synthesis, biochemical studies, cellular characterization, and structure-based computational studies of small molecules targeting the urokinase receptor, *Bio. Med. Chem.* 20, 4760–4773

[46] Karplus, M. (2002) Molecular dynamics simulations of biomolecules. *Acc Chem Res.* 35, 321–323

[47] Raha, K. et al. (2007) The role of quantum mechanics in structure-based drug design. *Drug Discov Today* 12, 725–731

[48] Peters, M.B. et al. (2006) Quantum mechanics in structure-based drug design. *Curr. Opin. Drug Discov. Devel.* 9, 370–379

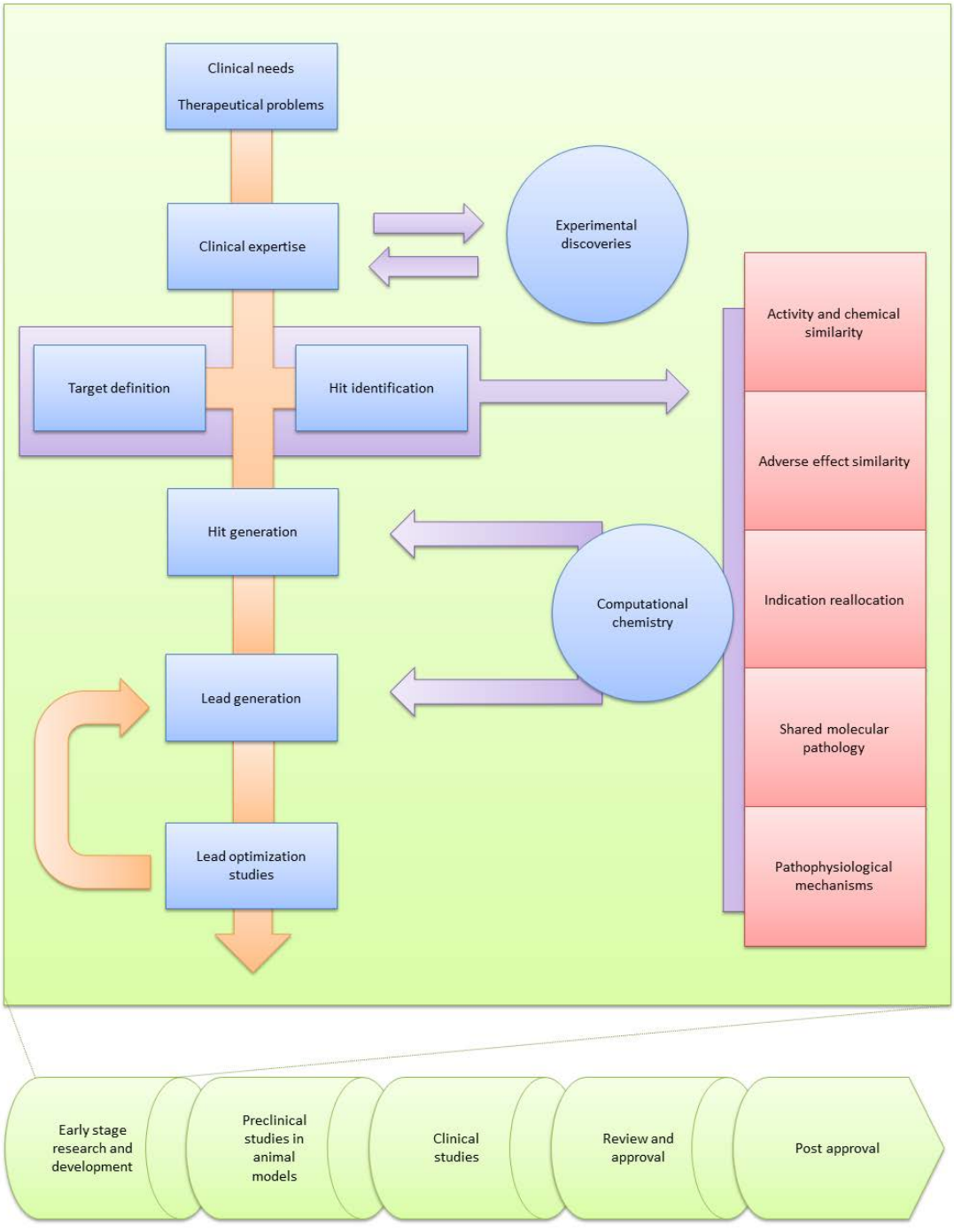
[49] Gleeson, M.P. et al. (2011) In-silico ADME models: a general assessment of their utility in drug discovery applications. *Curr. Top. Med. Chem.* 11, 358–381

[50] Liao, C. et al. (2011) Software and resources for computational medicinal chemistry. *Future Med. Chem.* 3, 1057–1085

[51] Vogt, M. and Bajorath, J. (2012) Chemoinformatics: A view of the field and current trends in method development, *Bioorg. Med. Chem.* 20, 5317–5323

Figure 1

Proposed translational model of drug development. The global process of drug development, with the stages in which is applied the translational approach between clinical practice and computational chemistry.



Box 1

Auxiliary computational techniques for drug design

Molecular mechanics and dynamics software

Molecular dynamics simulations are based on Newton's equations of motion. Molecular dynamics is very useful for understanding the dynamic behavior of proteins or other biological macromolecules, from fast internal motions to slow conformational changes or even protein-folding processes. These simulations incorporate flexibility of both the receptor and the ligand, coming closer to the ideal of induced fit by enhanced complementarity and interaction Molecular dynamics simulations integrate explicit solvent molecules, creating a more mimetic environment of the biological conditions, adding the solvent's effect on the stability of the ligand–protein complexes [46]. Thus, the results from MD simulations can be employed as target for docking studies or the technique can be employed to refine docked complexes [33].

Quantum mechanics software

Being the nuclei held together by electron orbitals governed by the laws of quantum mechanics, ligand-based and structure-based methods can be addressed using quantum mechanics methods. This fact as become reality due to the increase of central processing unit performance and the improvement of algorithms and software [47]. Quantum mechanics methods can be used to model small to medium-sized molecules, radicals and estimate activation energies for chemical and enzymatic reactions. The applications

in drug design include calculation of energies and optimization of structures of ligands and protein–ligand complexes, calculation of atomic point charges applicable to correcting the binding mode of a ligand obtained from docking studies, calculation of free binding energies and build of QSAR models [48].

ADME/T software

ADME/T prediction software is capable of predicting potential risks in pharmacokinetics and toxicology, with great benefit in the design of molecules that not only potentially interact with the putative receptor selected but also accomplish the criteria for being used as a drug in a safe dosage and posology. The concept consists in the development of statistical models supported by QSAR. The relationships established are not determined for prediction of activity over a receptor involved in a disease but to predict ADME/T features [49].

Molecular visualization software

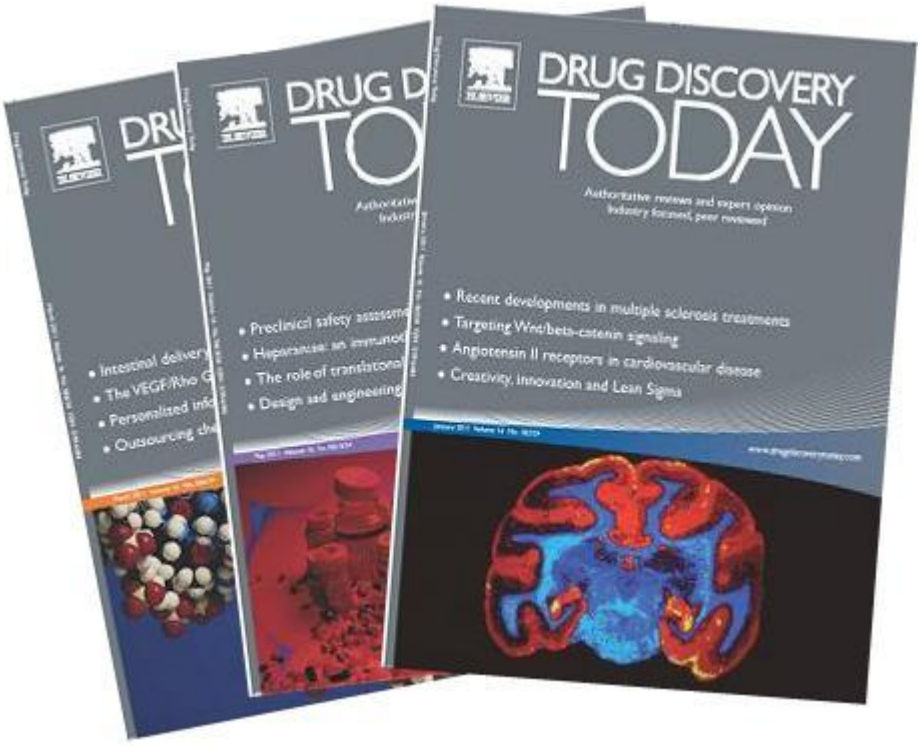
Molecular visualization programs are graphical user interfaces, through which the users can visualize and analyze their models and results, and can generate graphics for publications or reports. There is the possibility of analysis of density maps, supramolecular assemblies, sequence alignments, docking results, trajectories and conformational ensembles [50].

Chemoinformatics software

Chemoinformatics software consists in computational tools that assist in the acquirement, analysis and management of data of chemical compounds and their properties. The programs used prioritize on the management of information. Such requirements were frequently regarded as barriers by researchers, as the interchange of data between different programs usually requires some programming experience. The advent of visual workflow/ data pipelining environments diminished the problem at some extent. These computational environments provide the ability to graphically layout or build protocols and workflows, which can be reused, extended or rerun later also by other users [51].

Drug Discovery Today

INSTRUCTIONS FOR AUTHORS



ELSEVIER

DRUG DISCOVERY TODAY

Drug Discovery Today article types	Length
Guest Editorial articles provide a forum for a personal perspective on contemporary issues and controversies – something you are passionate about or that you think our readers will find thought-provoking.	• Max. 1500 words • Max. 10 references • 1 portrait photograph of corresponding author
Short Reviews cover fast-moving recent research topics and form the backbone of *Drug Discovery Today* content. There are three categories of short reviews: **Gene to Screen:** reviews dealing with the earliest part of drug discovery and all aspects involved in target identification, validation and assay development, particularly advances in genomics and proteomic technologies. **Informatics:** reviews focusing on the latest developments and advances in computational drug discovery. **Post Screen:** reviews covering the science from hit identification to candidate selection, clinical trial design and patent issues as well as associated technologies and business strategies. **Short Reviews** should provide a brief overview of the background and then concentrate on setting recent findings (past 1–2 years) in context. Although they may often tackle controversial topics, they must give a balanced view of developments and authors must never concentrate unduly on their own research. However, they should allow room for some speculation and debate. Articles should be preceded by an abstract of ~100 words. Additional explanatory text or definitions of specialist terms can be put in boxes. The Editor reserves the right to request that reviews be shortened or lengthened.	• Max. 3500 words • Max. 60 references • Max. 4 figures, boxes or tables • ~100-word abstract • 25–30-word teaser
Keynote Reviews cover topics in a broad, authoritative, timely and comprehensive manner and are longer than the traditional Short Reviews. They should provide a critical and comprehensive overview of the field, discussing important background information, key concepts and summaries of latest developments. Authors must give a balanced view of developments and must never concentrate unduly on their own research. Articles should be preceded by an abstract of ~100 words and a short (~100 words) author biography and photo. Authors are encouraged to include more figures, boxes and tables than might traditionally be found in Short Reviews. In addition, a glossary containing definitions of specialist terms and a box containing key resources (i.e. books, URLs) relevant to the topic should be included	• 6000–7000 words • Max. 100 references • Max. 7 figures, boxes or tables • ~100-word abstract • 25–30-word teaser • ~100-word author biogs (up to three biogs only) • Author photographs (up to three photos only

Foundation Reviews provide an introduction and comprehensive review of fundamental principles in drug discovery, forming an indispensable educational resource. These articles introduce beginners in the field to concepts and methods used in drug discovery whilst also providing a valuable tool for more experienced workers. These articles should be written in an accessible manner and provide important background information, a review of theory, application and methods in drug discovery, as well as a discussion of historical perspectives and future challenges. Authors are encouraged to include explanatory figures, boxes and tables. In addition, a glossary containing definitions of specialist terms and a box containing key resources (i.e. books, URLs) relevant to the topic should be included. Articles should be preceded by an abstract of ~100 words and a short (~100 words) author biography and photo.	• 6000–7000 words • Max. 100 references • Max. 7 figures, boxes or tables • ~100 word abstract • 25–30-word teaser • ~100-word author biogs (up to three biogs only) • Author photographs (up to three photos only)
Features are opinion articles on controversial topics or recent developments in the industry, or can cover strategic industry issues, new areas of research, profiles of new research organisations or industry trends. They should stimulate debate, cover controversial topics that are being hotly debated, present new models or hypotheses (along with suggestions for future experiments), or speculate on the meaning/interpretation of some new data. Articles that merely outline recent advances in a field rather than give an *opinion* on them are not suitable for this section.	• Max. 2000 words • Max. 30 references • Max. 3 figures
Biotech focus articles provide an overview of activities within the pharmaceutical biotechnology sector, focusing on recent developments, strategic issues or specific geographical hotspots. These articles can highlight advances and implications of new developments to this sector, can stimulate debate by covering controversial topics, or provide a forum for future perspectives and directions. Articles that focus on specific biotechnology regions should describe an overview of the biotechnology activities within that region. They should report the local research expertise, and provide readers with a general background of the different companies and their particular strengths. In addition, a short description of how the area developed (for example, as a result of spin-offs from a local university) would be of interest. Aerial photographs of the biotechnology park can be included for illustrative purposes, or regional maps.	• Max. 2500 words • Max. 4 figures, boxes or tables • Max. 5 references

Please submit completed manuscripts via our EES site

http://www.ees.elsevier.com/drudis

For specific instructions on how to submit via EES please see:
http://support.elsevier.com/app/answers/detail/a_id/353/c/6261

For queries on these instructions please contact our Editorial Office:
E-mail: DDT@drugdiscoverytoday.com

Please note that manuscripts that significantly exceed the stated length will be returned to the author.
Please provide current contact details for **3–4 suitable referees**. Although there is no guarantee that we will approach these specific individuals, the provision of contact details may help to ensure rapid progress through the peer-review process.
If you have any difficulties preparing your text or figures, please contact the editorial office for clarification

Ethics in Publishing

For information on Ethics in Publishing and Ethical guidelines for journal publication see http://www.elsevier.com/publishingethics and http://www.elsevier.com/ethicalguidelines.

Conflicts of Interest

All authors are requested to disclose any actual or potential conflict of interest including any financial, personal or other relationships with other people or organizations within three years of beginning the submitted work that could inappropriately influence, or be perceived to influence, their work. See also http://www.elsevier.com/conflictsofinterest.

Submission declaration

Submission of an article implies that the work described has not been published previously (except in the form of an abstract or as part of a published lecture or academic thesis), that it is not under consideration for publication elsewhere, that its publication is approved by all authors and tacitly or explicitly by the responsible authorities where the work was carried out, and that, if accepted, it will not be published elsewhere including electronically in the same form, in English or in any other language, without the written consent of the copyright-holder

Copyright

Upon acceptance of an article, authors will be asked to complete a 'Journal Publishing Agreement' (for more information on this and copyright see http://www.elsevier.com/copyright). Acceptance of the agreement will ensure the widest possible dissemination of information. An e-mail will be sent to the corresponding author confirming receipt of the manuscript together with a 'Journal Publishing Agreement' form or a link to the online version of this agreement.
Subscribers may reproduce tables of contents or prepare lists of articles including abstracts for internal circulation within their institutions. Permission of the Publisher is required for resale or distribution outside the institution and for all other derivative works, including compilations and translations (please consult http://www.elsevier.com/permissions). If excerpts from other copyrighted works are included, the author(s) must obtain written permission from the copyright owners and credit the source(s) in the article. Elsevier has preprinted forms for use by authors in these cases: please consult http://www.elsevier.com/permissions

Retained author rights

As an author you (or your employer or institution) retain certain rights; for details you are referred to: http://www.elsevier.com/authorsrights.

Role of the funding source

You are requested to identify who provided financial support for the conduct of the research and/or preparation of the article and to briefly describe the role of the sponsor(s), if any, in study design; in the collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the paper for publication. If the funding source(s) had no such involvement then this should be stated. Please see http://www.elsevier.com/funding.

Funding body agreements and policies

Elsevier has established agreements and developed policies to allow authors whose articles appear in journals published by Elsevier, to comply with potential manuscript archiving requirements as specified as conditions of their grant awards. To learn more about existing agreements and policies please visit http://www.elsevier.com/fundingbodies.

Language and language services

Please write your text in good English (American or British usage is accepted, but not a mixture of these). Authors who require information about language editing and copyediting services pre- and post-submission please visit http://www.elsevier.com/languageediting or our customer support site at http://epsupport.elsevier.com for more information.

Submission

Submission to this journal proceeds totally online and you will be guided stepwise through the creation and uploading of your files. The system automatically converts source files to a single PDF file of the article, which is used in the peer-review process. Please note that even though manuscript source files are converted to PDF files at submission for the review process, these source files are needed for further processing after acceptance. **Please do NOT submit .PDF files**. All correspondence, including notification of the Editor's decision and requests for revision, takes place by e-mail, removing the need for a paper trail.

Use of wordprocessing software

It is important that the file be saved in the native format of the wordprocessor used. The text should be in single-column format. Keep the layout of the text as simple as possible. Most formatting codes will be removed and replaced on processing the article. In particular, do not use the wordprocessor's options to justify text or to hyphenate words. However, do use bold face, italics, subscripts, superscripts etc. Do not embed "graphically designed" equations or tables, but prepare these using the wordprocessor's facility. When preparing tables, if you are using a table grid, use only one grid for each individual table and not a grid for each row. If no grid is used, use tabs, not spaces, to align columns. The electronic text should be prepared in a way very similar to that of conventional manuscripts (see also the Guide to Publishing with Elsevier: http://www.elsevier.com/guidepublication). Do not import the figures into the text file but, instead, indicate their approximate locations directly in the electronic text and on the manuscript. See also the section on Electronic illustrations.

To avoid unnecessary errors you are strongly advised to use the "spell-check" and "grammar-check" functions of your wordprocessor.

Graphical abstract

A Graphical abstract is optional and should summarize the contents of the paper in a concise, pictorial form designed to capture the attention of a wide readership online. Authors must provide images that clearly represent the work described in the paper. Graphical abstracts should be submitted as a separate file in the online submission system. Maximum image size: 400 × 600 pixels (h × w, recommended size 200 × 500 pixels). Preferred file types: TIFF, EPS, PDF or MS Office files. See http://www.elsevier.com/graphicalabstracts for examples.

Research highlights

Research highlights are mandatory for this journal. They consist of a short collection of bullet points that convey the core findings of the article and should be submitted in a separate file in the online submission system. Please use 'Research highlights' in the file name and include 3 to 5 bullet points (maximum 85 characters per bullet point including spaces). See http://www.elsevier.com/researchhighlights for examples.

Video data

Elsevier accepts video material and animation sequences to support and enhance your scientific research. Authors who have video or animation files that they wish to submit with their article are strongly encouraged to include these within the body of the article. This can be done in the same way as a figure or table by referring to the video or animation content and noting in the body text where it should be placed. All submitted files should be properly labeled so that they directly relate to the video file's content. In order to ensure that your video or animation material is directly usable, please provide the files in one of our recommended file formats with a maximum size of 10 MB. Video and animation files supplied will be published online in the electronic version of your article in Elsevier Web products, including ScienceDirect: http://www.sciencedirect.com. Please supply 'stills' with your files: you can choose any frame from the video or animation or make a separate image. These will be used instead of standard icons and will personalize the link to your video data. For more detailed instructions please visit our video instruction pages at http://www.elsevier.com/artworkinstructions. Note: since video and animation cannot be embedded in the print version of the journal, please provide text for both the electronic and the print version for the portions of the article that refer to this content.

Supplementary data

Elsevier accepts electronic supplementary material to support and enhance your scientific research. Supplementary files offer the author additional possibilities to publish supporting applications, high-resolution images, background datasets, sound clips and more. Supplementary files supplied will be published online alongside the electronic version of your article in Elsevier Web products, including ScienceDirect: http://www.sciencedirect.com. In order to ensure that your submitted material is directly usable, please provide the data in one of our recommended file formats. Authors should submit the material in electronic format together with the article and supply a concise and descriptive caption for each file. For more detailed instructions please visit our artwork instruction pages at http://www.elsevier.com/artworkinstructions.

Checklist for Authors: *Drug Discovery Today*

(Please tick the boxes once the following have been included in your manuscript)

- **Before you begin submission**

☐ Source files for the manuscript, figures and tables

☐ File containing 3 to 5 research highlights

☐ Names and e-mail addresses of 3 or 4 referees.

☐ Optional: graphical abstract, video abstract and supplementary data files

• Title page (page 1)

☐ Short title (<8 words long, enticing, relevant to the content).

☐ Authors' names (no more than 5 names, first names and surnames in full, with middle initials).

☐ Authors' current addresses.

☐ One corresponding e-mail address written as: *Corresponding author*: Smith, A.B. (absmith@ucl.ac.uk).

☐ Keywords (6 maximum) for indexing purposes.

☐ A teaser (1 sentence, 25–30 words maximum) to convey to the reader why the article is relevant and interesting.

• Main text (page 2)

For details on word count, please see article type table.

The word count represents the number of words in the body of the text and excludes the abstract, figure legends, text boxes and reference list.

☐ Abstract: all Review articles should be prefaced by an abstract. This should attract the reader's interest and contain sufficient information for the reader to be able to appreciate the relevance of the article. It should include a brief description of the topic, background information necessary to appreciate the importance of the discussion, and bold statements of the main conclusions or predictions, rather than promises that a particular subject 'will be discussed'. References should not be included and abbreviations avoided.

☐ Introductory section; no heading required

☐ Subheadings: 4–6 short descriptive subheadings should be used to break up the text into logical sections. Second-level subheadings can be used if necessary.

☐ Conclusion: the article should finish with a paragraph emphasizing the prospects for future research as well as summarizing the current state of knowledge.

☐ Citing references: please use numbers in square brackets, in order of citation: e.g. [1] [2,3] [4–7], not alphabetical. If tables and figure legends contain any refs in addition to those cited in the main text, main text refs must be numbered first, followed by additional refs in the tables, followed by additional refs in the figure legends.

☐ Acknowledgements: placed before reference list, should not acknowledge grants or original research contributors.

☐ Algebra: numerical variables in italic; categories and groups in roman; vectors in bold

☐ Define all abbreviations on first mention

• Reference lists

How many references? Please see Article type table for number of references allowed.

☐ Unpublished work, PhD theses and URLs/website addresses must be cited in main text, not in reference lists.

☐ Unpublished work: cited in main text in parentheses as: (Q. Cumber-Patch *et al.*, unpublished).

☐ PhD theses: cited in main text in parentheses: (R. Arthur Goode, PhD thesis, University of Hawaii, 1988).

☐ URLs/website addresses: cited in main text in parentheses: (see: http://www.xxx.yyy.zzz).

☐ References in main text, boxes and figures are numbered, and listed at the end of the main text.

☐ In tables, references should be cited in numbers, in a separate column, and listed at the end of the main text.

☐ References listed in order of citation, not alphabetically, with one reference per number.

☐ **For journal references**: please give authors' names (if two authors, print both names separated by 'and'; if three or more authors, use *et al*. after first author); date (in parentheses); title (in roman text); abbreviate journal name using *Biological Abstracts*; volume; and complete page range. For example:

1 Gold, B. (2002) Effect of cationic charge localization on DNA structure. *Biopolymers* 65, 173–179
2 Han, Y. and Barillas-Mury, C. (2002) Implications of Time Bomb model of ookinete invasion of midgut cells. *Insect Biochem. Mol. Biol.* 32, 1311
3 Gruber, D.M. *et al.* (1999) Progesterone and neurology. *Gynecol. Endocrinol.* 4, 41–45
4 Jovani, R. Malaria transmission, sex ratio, and erythrocytes with two gametocytes. *Trends Parasitol.* (in press)

☐ **For online journal references**: please give authors' names (as above); date (in parentheses); title (in roman text); abbreviate journal name using *Biological Abstracts*; the digital object identifier (DOI) number; and the website of the journal. For example:

5 Jiang, J.C. *et al.* (2000) An intervention resembling caloric restriction prolongs life span and retards aging in yeast. *FASEB J.* DOI: 10.1096/fj.00-242fje (http://www.fasebj.org)

☐ **For book references:**
For whole books: please give editors' names; date (in parentheses); title (in *italics*); and publisher. For example:
1 Chowdhury, N. and Alonso Aguirre, A., eds (2001) *Helminths of Wildlife*, Science Publishers Inc.
For book chapters: please give chapter authors; date (in parentheses); chapter title; book title (in *italics*); editors' names; page numbers and publisher. For example:
35 Clutton-Brock, T. and Godfray, H.C.J. (1991) Parental investment. In *Behavioural Ecology* (3rd edn) (Krebs, J.R. and Davies, N.B., eds), pp. 234–262, Blackwell

☐ **For patent references**:

23 Bloggs, J. *et al.* Company name that actually owns the patent. Title of patent, Code

• Accession numbers

Accession numbers are unique identifiers in bioinformatics allocated to nucleotide and protein sequences to allow tracking of different versions of that sequence record and the associated sequence in a data repository [e.g., databases at the National Center for Biotechnical Information (NCBI) at the National Library of Medicine ('GenBank') and the Worldwide Protein Data Bank]. There are different types of accession numbers in use based on the type of sequence cited, each of which uses a different coding. Authors should explicitly mention the *type of accession number together with the actual number*, bearing in mind that an error in a letter or number can result in a dead link in the online version of the article. Please use the following format: accession number type ID: xxxx (e.g., MMDB ID: 12345; PDB ID: 1TUP). Note that in the final version of the *electronic copy*, accession numbers will be linked to the appropriate database, enabling readers to go directly to that source from the article

Possible ways of citing accession numbers in the text:

1 Sequences for introns 1 and 21 (NH0349G04: accession number AC008172.1) sequenced in GenBank.

2 The accession numbers for each sequence are as follows: BAA78620 (Amphihox1), P09022 (mouse Hox-A1) and CAB57787 (*Drosophila* Lab).

• Additional material (Boxes, Tables and Figures)

Please see Article Type Table for the number of separate pieces of additional material allowed.

Boxes

Boxes can be used for additional explanatory material, which, although essential, interrupts the flow of the text (e.g. mathematical models, glossaries, methodologies and historical notes). Can contain figures and tables. Should be <500 words long.

☐ Have you cited all boxes in the main text?

☐ Please provide a single-sentence title for the box (<8 words), double-space box text (500 words max.).

☐ Explain all abbreviations at first mention unless already defined in main text.

Tables

☐ Have you cited all tables in the text?

☐ Please provide single-sentence title for the table, double-space and run-on all text.

☐ Footnotes: help the reader to understand the table without referring to the main text. Use superscript letters[a,b] to refer to footnotes in alphabetical order. All abbreviations, symbols etc. must be explained in a footnote unless abbreviations have been previously defined in the main text.

☐ References cited in tables should be in a separate column and listed in the main reference list (in sequence from end of main reference list).

Figures

- ☐ Have you cited all figures in the main text and/or box text?
- ☐ Have you obtained permission to reproduce copyrighted material (i.e. material, such as figures, tables or excerpts, that has already been published elsewhere) from the copyright owners of that material
- ☐ Have you acknowledged, in the figure legend, the original source of previously published material?
- ☐ Please supply individual, editable files of each of your figures. These files should be in the format in which they were originally created, rather than imported into other programs.
- ☐ Figure labels: always first letter capital, then remainder lower-case (not bold or italic, except for species).
- ☐ Please provide a figure legend to help the reader to understand the figure without referring to the main text, including: a short title; scale bar (if appropriate); references (should be listed in the main reference list, in sequence from end of list); and explain all abbreviations, symbols and colour codes etc. Please place figure legends at the end of main text (after reference list) and not next to the figure. Figure legends should concisely describe what is shown in the figure, and should allow the figure to stand alone without reference to the text.
- ☐ All abbreviations used in figure are explained in legend.

For figure submission guidelines please see:

http://www.elsevier.com/artworkinstructions