Taylor & Francis
Taylor & Francis Group

Expert Opinion on Drug Discovery

ISSN: 1746-0441 (Print) 1746-045X (Online) Journal homepage: http://www.tandfonline.com/loi/iedc20

The application of molecular topology for ulcerative colitis drug discovery

Carolina L. Bellera, Mauricio E. Di Ianni & Alan Talevi

To cite this article: Carolina L. Bellera, Mauricio E. Di Ianni & Alan Talevi (2017): The application of molecular topology for ulcerative colitis drug discovery, Expert Opinion on Drug Discovery, DOI: 10.1080/17460441.2018.1396314

To link to this article: http://dx.doi.org/10.1080/17460441.2018.1396314

Published online: 01 Nov 2017.

EXPERT OPINION ON DRUG DISCOVERY, 2018
https://doi.org/10.1080/17460441.2018.1396314

Taylor & Francis
Taylor & Francis Group

REVIEW

The application of molecular topology for ulcerative colitis drug discovery

Carolina L. Bellera, Mauricio E. Di Ianni and Alan Talevi

Medicinal Chemistry/Laboratory of Bioactive Research and Development, Department of Biological Sciences, Faculty of Exact Sciences, University of La Plata (UNLP), La Plata, Buenos Aires, Argentina

ABSTRACT

Introduction: Although the therapeutic arsenal against ulcerative colitis has greatly expanded (including the revolutionary advent of biologics), there remain patients who are refractory to current medications while the safety of the available therapeutics could also be improved. Molecular topology provides a theoretic framework for the discovery of new therapeutic agents in a very efficient manner, and its applications in the field of ulcerative colitis have slowly begun to flourish.

Areas covered: After discussing the basics of molecular topology, the authors review QSAR models focusing on validated targets for the treatment of ulcerative colitis, entirely or partially based on topological descriptors.

Expert opinion: The application of molecular topology to ulcerative colitis drug discovery is still very limited, and many of the existing reports seem to be strictly theoretic, with no experimental validation or practical applications. Interestingly, mechanism-independent models based on phenotypic responses have recently been reported. Such models are in agreement with the recent interest raised by network pharmacology as a potential solution for complex disorders. These and other similar studies applying molecular topology suggest that some therapeutic categories may present a 'topological pattern' that goes beyond a specific mechanism of action.

ARTICLE HISTORY
Received 2 August 2017
Accepted 20 October 2017

KEYWORDS
Ulcerative colitis; inflammatory bowel disease; drug discovery; molecular topology; topological indices; molecular descriptors; QSAR; virtual screening

1. Introduction

Ulcerative colitis (UC) is a type of inflammatory bowel disease. It is a chronic condition that affects the colonic mucosa, starting in the rectum and extending to part of, or the entire, colon [1]. The disease is frequently characterized by alternating cycles of exacerbation and remission. Patients with UC suffer from rectal bleeding and bloody stools, diarrhea, urgency, and abdominal pain. It can be regarded as a complex disorder, with environmental [2–4] and constitutive [4,5] components. The complex and not fully understood physiopathology comprises the interplay between the dysregulated immune system, the intestinal flora, and the epithelial barrier function. A summary of the immunopathogenic mechanisms known to be involved in UC is presented in Figure 1. The diagram has been produced from information in the works of Danese et al. [6], Hindryckx et al. [7], and Fuss and Strober [8].

There exists a wide array of currently approved therapies for the treatment of UC, comprising a diversity of mechanisms of action and both small molecules and biologics. The choice of the therapeutic approach depends on a number of factors such as disease severity, treatment goal (to induce or maintain remission), response to treatment, and tolerability [9,10].

The first-line therapy for mild-to-moderate UC is based on administration of aminosalicylates, in particular, 5-aminosalycilic acid (5-ASA) or 5-ASA prodrugs (e.g. balsalazide, sulfasalazine). Their mechanisms of action are diverse and not entirely understood. 5-ASA elicits a potent inhibitory effect on a number of pro-inflammatory mediators, such as leukotrienes, interleukin-1 (IL-1), and tumor necrosis factor alpha (TNF-α) [11]. It has been observed that 5-ASA is an agonist for the peroxisome proliferator activated receptor gamma, which plays a key role in the regulation of inflammatory signaling pathways [12,13]. While rectal administration of 5-ASA or its prodrugs allows effective treatment of distal UC, acceptance of such route varies significantly from country to country, which relates to psychosocial issues but also, in some cases, medical staff attitudes, and beliefs [14]. In order to exploit the less invasive oral route of administration, delayed-release dosage forms have been designed that allow building up effective levels of the drug in the colon simultaneously alleviating the high burden of daily dosage units that would be required otherwise.

Glucocorticosteroids are the first choice for patients refractory to 5-ASA, and to induce remission in moderate-to-severe UC, though they seem to have no benefit as maintenance therapy [7,9,15]. What is more, they present risk of steroid dependency [16], steroid-induced metabolic disturbances [15], and opportunistic infections [17]. The oral or rectal routes are preferred for mild-to-moderate cases; second-generation corticosteroids as budesonide and beclomethasone dipropionate display high first-pass effect and thus reduced systemic activity [10,15]. Intravenous corticosteroid rescue therapy is also used for acute severe UC cases.

Immunomodulators (e.g. thiopurines) are administered to patients with moderate UC who are refractory or do not tolerate steroids. Although they have shown efficacy for both

CONTACT Alan Talevi atalevi@biol.unlp.edu.ar

Article highlights

- Currently approved treatments against ulcerative colitis have been discovered through immunological studies. The number of studies addressing the discovery of drug candidates using computer-aided methodologies is still scarce. Molecular Topology might be the dominant approach for ligand-based approximations in the field of ulcerative colitis computer-aided drug discovery.
- Molecular Topology, jointly with QSAR, can be applied to explore the ever growing chemical space in a very efficient manner.
- Topological indices are graph-invariant descriptors obtained from chemical graphs by algebraic operations. They are, in general, conformation- and orientation-independent (thus, no geometry optimization of any sort is required to obtain their values). Generally speaking, they are insensitive to geometric and space isomers.
- Applications of Molecular Topology in different fields of drug discovery suggest that, independently of their particular mechanisms of action, drugs that belong to a given therapeutic category tend to share similar topologies.
- Among the many QSAR studies that use molecular topology to identify novel drug candidates for ulcerative colitis, a considerable number of articles report models that include topological autocorrelations weighted by different atomic properties (e.g. electronegativity, Van der Waals volumes). Interestingly, those descriptors are strongly related with the pharmacophore concept, but from a topological perspective instead of a space geometric one.
- Being a complex disease, ulcerative colitis should be approached from network based pharmacology perspective. QSAR models that model phenotypic responses (e.g. activity in an *in vivo* model of colitis) could prove very useful to find novel effective treatments.

This box summarizes key points contained in the article.

induction and maintenance of remission [10], they are preferably used for the later [9,15]. They present a number of deleterious effects (in around 20% of the patients) that limit their application, including bone marrow suppression, pancreatitis, hepatotoxicity, and increase risk of opportunistic infections (synergic with that of steroids), among others [9,10,15].

A diversity of biologics targeting specific immunological pathways comprise the last generation UC treatments [17], including anti-TNF agents (infliximab, adalimumab, golimumab) and anti-adhesion therapies (vedolizumab), though other emerging therapies are presently at clinical trials. Biologics are used in moderate-to-severe UC, and in patients that have shown inadequate response or that have medical contraindications to conventional therapies. A drawback of biologics used to treat UC is their risk of severe adverse events, such as acute infusion reaction, severe serum sickness, opportunistic infections, or lymphoma [7,9,18,19], although at least some of these complications appear to be solved or reduced with the most recently approved medications. In some cases (anti-TNF agents), diminished efficacy after repeated treatment has been observed [10,18]. Other aspects to be considered are that biologics must be administered parenterally and that they are significantly more expensive than conventional therapies [20], which may limit their accessibility in some scenarios.

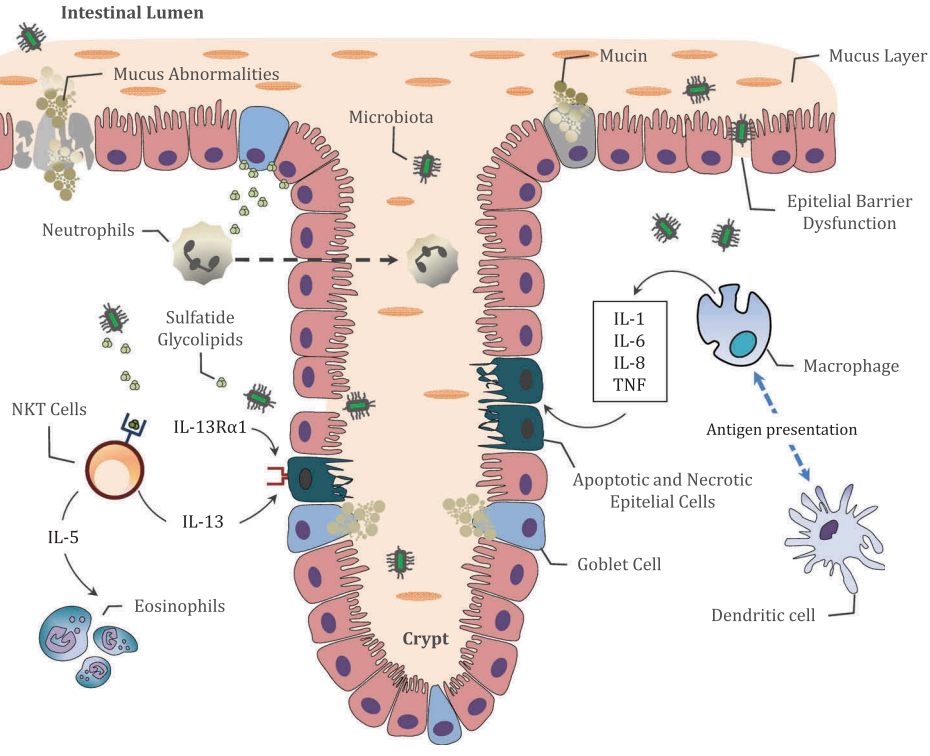


Figure 1. Simplified diagram of the immunopathogenic mechanisms involved in UC. Mucosal barrier abnormalities in UC may be developed by many issues, namely: genetic abnormalities resulting in the reduced production of mucus components and thus compromise to the initial barrier to microbial access to epithelial cells; release of glycolipids from epithelial cells and bacteria and subsequent stimulation of NK T cells that cause epithelial cell barrier abnormalities via direct cytotoxicity or production of IL-13, which in turn leads to enhanced absorption of bacterial products and the generation of antibacterial antibodies; infiltration of neutrophils between epithelial cells and into the gut lumen; tight junction defects. Increased absorption of bacterial products stimulates dendritic cells and macrophages, resulting in the production of proinflammatory cytokines.

Whereas great advance has been made in the last decade regarding UC management, there are still research goals to be met in the development of novel drugs and dosage forms [9,21], among them improved efficacy on refractory patients, improved safety and development of formulations capable of delivering effective levels of the therapeutic agent in the colon with reduced systemic absorption (in particular, for local treatments).

Molecular topology (MT) is an area of mathematical chemistry consisting of the topological description of molecular structures [22–24]. Starting from mathematical graph theory and integrating it with chemical notions, MT studies the connectivity of the atoms forming a molecule. From appropriate graph representations, MT most frequently derives matrixes and, by algebraically operating on such matrixes, yields a great diversity of molecular descriptors (topological indices or topological descriptors) that can be used both in the frame of Quantitative Structure-Activity Relationships (QSAR) theory and for descriptive statistical purposes (i.e. to rationally explore the chemical space). By far, the most common application of topological descriptors involves finding a quantitative correlation between the molecular structure and one or more biological responses (expressed as a continue output or as a discrete class label output). In spite of years of strong dominance by the structure-based methods in the drug discovery arena, the QSAR field seems healthy and even expanding [25,26]. The use of MT in QSAR (either as purely topological models or by combining topological indexes with other sort of molecular descriptors) has exponentially grown in the later years: today, MT covers more than 20% of all articles dealing with QSAR [22]. As we will review in the following sections, it is the dominant approach for ligand-based approximations in the field of UC drug discovery [27–30].

Here, we will first present a brief introduction to MT, after which we will discuss the existing reports on MT applications for UC drug discovery, either reporting 'pure' topological models or combinations of topological indexes with other classes of molecular descriptors. The authors would like to state that they are not experts on UC, but have years of experience on ligand-based drug discovery approximations and MT applications. Accordingly, the article is focused on the computational aspects of the reviewed studies.

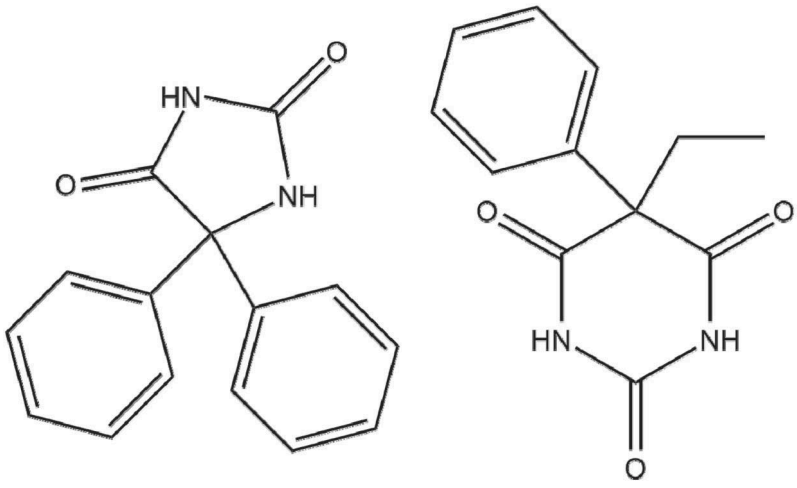
2. Virtual screening, QSAR, and molecular topology

The accessible chemical space is growing fast and exponentially. Back in 2012, the number of compounds indexed in the Chemical Abstract Service accounted to about 70 million. Today, that number totals more than 130 million, which speaks of an astonishing average expansion of 12 million new chemical entities per year during the last 5-year period. Judging from the current numbers in PubChem Compound (at present more than 93 million entries) and ZINC database (today over 35 million entries), a quite considerable proportion of the accessible chemical space comprises small, drug-like molecules. This portrays an encouraging picture, since it speaks of an impressive chemical diversity to seek for new lead compounds which in turn will be submitted to lead optimization programs. However, it also poses a valid question: can we explore such a vast chemical space in an efficient manner?

Efficient screening tools have been developed to address this challenge, including wet approaches such as high-throughput screening and ultrahigh-throughput screening, which couple miniaturization and automated screening platforms [31,32]. Alternatively, *in silico* tools, i.e. virtual screening (VS), can be applied. In VS campaigns, a wide spectrum of computational methods is used to rank digital chemical collections or libraries to establish which compounds (the top-ranked ones) are more likely to obtain positive results when subjected to *in vitro* and/or *in vivo* models. VS has been conceived to minimize the volume of experimental testing and optimize the results at such stage, thus being advantageous in terms of cost-efficiency, bioethics, and environmental impact. Some versions of it can be performed with very accessible technology, especially ligand-based approximations, with many valuable resources to implement VS campaigns (from specialized software to online chemical repositories) being freely available.

Whereas the advent of high-performance computing and low-cost parallel computing has allowed the use of complex, computational demanding approaches such as structure-based approximations to explore large chemical databases [33,34], it is debatable if the increment in computing power will cope with the exponential, continuing expansion of the chemical space to make possible large-scale exploration of the known small molecule space (or, furthermore, biologics). Possibly, efficient ligand-based approaches will persist as independent approaches, and also as initial screening or prescreening tools to reduce the number of candidates submitted to structure-based methods to a manageable size. Undoubtedly, among the most efficient ligand-based approximations to explore large chemical collections are two-dimensional (2D) similarity-based methods [35,36] and MT-based QSAR [22,24]. An additional advantage of both 2D similarity and MT is that, being conformation-independent approaches, they require little or no molecular pre-processing (besides inevitable chemical structures curation). This feature renders these methods highly reproducible and less dependent on user choices. VS is the most frequent application of MT, but it is also possible to approach drug design through the inverse QSAR paradigm [22,23,37,38] (designing novel compounds with the topological requirements needed to elicit a given biological activity).

Interestingly, 2D similarity searches and MT are extensively related, especially if we take into consideration that circular molecular fingerprints (for example, extended connectivity fingerprints) are being increasingly applied in chemical search engines and similarity-based VS. In fact, the top-ranked compounds in most similarity searches are compounds with a very close topology to the query molecules. It has been stated that the discovery of phenytoin, more than a century ago, was the result of what we would now call a 2D similarity method [39]: phenytoin was the product of a search among nonsedative structural relatives of phenobarbital for agents capable of suppressing electroshock convulsions in laboratory animals (Figure 2). There are many successful stories regarding VS campaigns based on MT that have helped prioritizing compounds to the preclinical stage of development. For instance, Talevi et al. used a topological model to screen the Merck Index 13th database and discovered the anticonvulsant effects



phenytoin phenobarbital

Figure 2. Chemical structures of anticonvulsants phenytoin and phenobarbital.

of propylparaben and methylparaben, which were presumed to act by blocking voltage-operated sodium channels [40]. The blocking effects of propylparaben on sodium channels was later confirmed [41], which could explain its effects against the epileptiform activity induced by 4-aminopyridine in hipoccampal CA1 pyramidal cells [42]. Moreover, propylparaben in combination with diazepam showed neuroprotective effects in the pilocarpine-induced status epilepticus model in rats [43]. In another successful MT application, Bellera and co-workers used a topological model in a drug repurposing campaign and found that the antibiotic clofazimine and the antihypertensive benidipine display trypanocidal effects against *Trypanosoma cruzi* (etiologic agent of Chagas disease). The trypanocidal effects of these drugs were later confirmed in acute and chronic mice models of Chagas [44,45]. Besides hit identification, MT has found many other interesting applications within the drug discovery field: for instance, it has been applied to explain the differences between approved drugs, clinical candidates, and bioactive compounds [46] and even to estimate the safe starting dose in phase I clinical trials [47].

But what is MT? MT is intrinsically related to mathematical graph theory [48]. Graphs are mathematical structures used to describe pairwise relations between objects. A graph is made up of *vertices* or *nodes*, which are connected by *edges*. In the realm of Chemistry, graph theory produces chemical graph theory. Atoms are thus represented as vertices whereas chemical bonds become edges. Such molecular graph (and graph-derived elements such as topological matrixes and topological indices) can prove valuable to describe the interconnectivity between atoms in a molecule and, therefore, its topology (e.g. linearity, branching, presence and distribution of rings). What is more, vertices and edges can be embedded with additional chemical information, expanding the descriptive power of graph-derived mathematical entities. This later possibility is extremely relevant to model bioactivities, since heteroatoms come to play in almost any ligand-target binding event.

At the core of the chemical graph theory lies the *adjacency matrix*. From a hydrogen-suppressed molecular representation, a $N \times N$ symmetric matrix can be obtained whose elements *Aij* equal 1 if vertices *i* and *j* are directly connected through a covalent chemical bond, and 0 otherwise. The sum of all entries in the *i*th row or the *j*th column provides the *degree* or *topological valence* δ of vertex *i* or *j*, respectively. Professor Milan Randić defined the first connectivity index (now known as Randić index) back in 1975 [49]. It was defined as the sum of the degrees of the two vertices adjacent to each edge, extended to all edges of the graph. Another very relevant topological matrix is the *distance matrix*, whose elements *Dij* equal the number of edges joining two vertices *i* and *j* by the shortest path, provided that *i and j* are different, or 0 otherwise. The first topological index ever defined, the Wiener index, equals one half of the sum of all entries in the distance matrix [50]. These two examples (the simpler at hand, illustrated in Figure 3) depict the general procedure to obtain topological descriptors. It is interesting to note that the numbering of the vertices of the graph does not influence the value of the graph-derived descriptors: they are *graph invariants*. In principle, the molecular graph is not influenced by any deformation introduced to the molecule: topological indices are conformation-independent unless deliberately pursued otherwise. Reproducibility and ease of calculation are thus two of the important (and interrelated) virtues of topological descriptors. No conformational analysis, no geometry optimization, no orientation, or conformation-related decisions are required to compute topological descriptor. The modeler is released from the burden of answering the difficult question 'What conformation should be used to compute a molecular descriptor?', and from the noise that defining a conformer could introduce to the QSAR model [51,52]. Note that the former question is particularly difficult to answer if, in the frame of a VS campaign, one pretends to apply a QSAR model to a large chemical database. At the other side of the coin, the values of the topological descriptors are usually insensitive to space or even geometry isomers, with some very specific exceptions (see, for instance, [53]).

Frequently, informative and more complex topological descriptors are derived through modified adjacency or distance matrixes. For instance, the valence topological charge indices

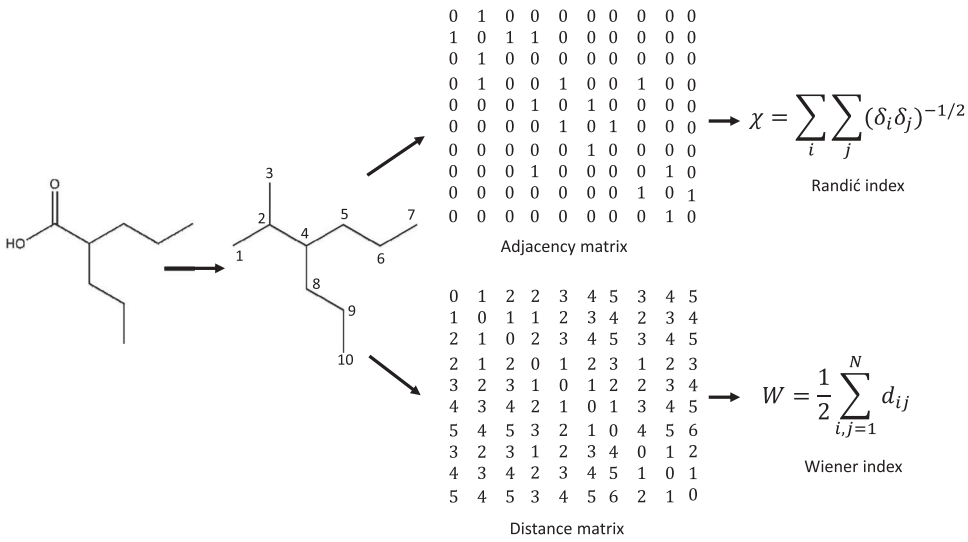


Figure 3. From a molecular representation, a hydrogen-depleted graph is obtained and its vertices arbitrarily numbered. The adjacency matrix (up) and the distance matrix (down) are obtained from such graph.

are obtained from the 'electronegativity-modified adjacency matrix,' which is similar to the adjacency matrix, but with the elements of the main diagonal replaced by the corresponding Pauling electronegativity value of the *i*th atom, weighted by two for each non-carbon, non-hydrogen atom [23]. Such modified matrix is multiplied by the inverse square distance matrix and only then the values of the descriptors are calculated.

The reader may probably capture, from the previous paragraphs, that chemical interpretation of topological descriptors can be a challenging task that requires a combination of chemical and mathematical skills. Such process could be even more complicated if one considers more complex topological descriptors such as eigenvalues of a topological matrix or spectral moments [22]. In contrast, understanding the chemical meaning of strictly physicochemical descriptors such as the log transformation of octanol–water partition coefficient (log P) or the polar surface area is more immediate for someone with a strong chemical background.

A comprehensive review of the existing topological descriptors is out of the scope of this article and almost impossible due to the large number of such indices reported in literature. Nevertheless, the curious reader is directed to the essential handbook on molecular descriptors by Todeschini and Consonni [54].

Once the topological descriptors have been calculated, statistical correlations can be explored between a subset of descriptors (independent variables) from a descriptor pool, and the observed values of a biological response of interest (dependent variables). The QSAR approach will seek mathematical functions or algorithms that respond to the general from:

$$Y = f(x1, x2,\ x3, \ldots,\ xn)$$

where *Y* represents the dependent variable (model output or response) and *xi* represents the *i*th molecular descriptor allowed into the model. The response may well be a continuous response (in the case of regression models) or a discrete one or class label (in the case of classifiers). The form of the mathematical function correlating *Y* with a set of descriptors will depend on the modeling technique (e.g. it will be linear in multi-linear regression, and nonlinear in the case of neural networks). There are also a variety of feature selection techniques available (e.g. stepwise procedures, genetic algorithms, replacement method, among many others). Discussing the details of such approaches and the steps to be followed to build a QSAR model is out of scope. In general terms, however, the procedure comprises the following steps [22,51]: compilation of a data set of compounds for which the response of interest has been experimentally assessed; computing a pool of molecular descriptors; partitioning of the data set into representative training and test sets; using the training set to calibrate the model (selecting a subset of relevant descriptors from the pool and weighting their contribution to explain the variability in the response); validating the model (internally, externally and sometimes experimentally). Any application of the model will require careful applicability domain assessment [55,56], that is, estimating the response- and chemical structure-spaces where the model makes reliable predictions.

3. Molecular targets in UC

In the following section, we will discuss MT applications to the discovery of potential new treatments for UC. We have considered reports on QSAR models purely based on topological descriptors, but also studies that combined topological descriptors with non-topological ones.

We have reviewed those (rather few) reports specifically addressing the discovery of novel treatments for UC through MT, but also reports that, using the QSAR approximation jointly with topological descriptors, sought to identify compounds acting on validated targets for UC (even if the authors of the later were not seeking UC treatments specifically).

To identify validated targets of UC, we searched the Therapeutic Target Database (TTD) [57]. Twenty-two different targets were found whose activity, according to TTD records, is

Table 1. List of targets of UC drugs retrieved from TTD (a drug targeting each target is included as example).

Target	Drug	Status
Alpha 4 beta 7 integrin	Vedolizumab	Approved
Prostaglandin G/H synthase 1	Mezalazine	Approved
Oxidoreductase	Olsalazine	Approved
Tumor necrosis factor	Infliximab	Approved
Alkaline phosphatase	Recombinant human alkaline phosphatase	Research target
Prostaglandin E2 receptor, EP4 subtype	KAG-308	Research target
Serine/threonine-protein kinase mTOR	P-2281	Research target
Gastrin/cholecystokinin type B receptor	S-0509	Phase II
Guanylate cyclase receptor	SP-333	Phase II
Melanocortin receptor	ASP-3291	Phase II
mRNA of Intercellular adhesion molecule-1	Alicaforsen	Phase II
Mucosal addressin cell adhesion molecule 1	PF-00547659	Phase II
Small inducible cytokine B10	Eldelumab	Phase II
92 kDa type IV collagenase	Andecaliximab	Phase II
Interleukin-13 receptor	Tralokinumab	Phase III
mRNA of Nuclear factor kappa B (TLR9 agonist)	Cobitolimod	Phase III
Guanylyl cyclase C	Guanilib	Phase III
5-lipoxigenase	Zileuton	Phase III
Antithrombin III	Sulodexide	Research target
Potassium-transporting ATPase alpha chain 1	S-pantoprazole	Phase III
IL-1 and IL-6	TAK-114	Phase III
Integrin beta 7	Etrolizumab	Phase III

regulated by known or potential UC treatments. Always according to TTD records, 4 of them are known targets of approved UC treatments, while 14 are targets of drug candidates that reached clinical trials. The information retrieved from TTD is summarized in Table 1. The research (alphanumeric) code provided by TTD was replaced by the common name of the drug whenever available (e.g. eldelumab instead of Anti-IP-10 or BMS-936557; andecaliximab instead of GS-5745). It is worth noticing that, while valuable, TTD seems to be incomplete. For example, tofacitinib, a new class of Janus kinase (JAK) inhibitor that blocks JAK1, JAK2 and JAK3, to modulate the signaling of a variety of IL receptors [18], has been omitted despite recent positive results in Phase III trials [58]. The IL-23/Th17 axis and IL-12/23 are also absent among TTD results [59–61]. Such missing targets have been added to our search.

For each identified target, we performed literature searches in Scopus including the target name plus 'QSAR' or 'molecular topology.' Additionally, we have included in the search all documents containing the expressions 'ulcerative colitis' or 'inflammatory bowel disease' and 'QSAR' or 'molecular topology.'

4. Molecular topology in ulcerative colitis drug discovery

4.1. Nuclear factor kappa B

Nuclear factor kappa B consists of a family of transcription factors involved in proinflammatory signaling pathways [62,63]. Upon activation by microbial products and proinflammatory cytokines such as IL-1 and TNF-α, nuclear factor kappa-light-chain-enhancer of activated B cells (NF-κB) exerts a regulatory role in the transcription of an extensive number of genes, including cytokines, chemokines, cell adhesion molecules, factors of the complement cascade, and acute phase proteins.

Back in 2011, Gálvez-Llompart et al. reported the first MT-based VS campaign specifically focused in the discovery of new treatments for UC [28]. They compiled a 121-compound data set including 67 known NF-κB inhibitors and 54 presumably inactive compounds retrieved from Merck Index, without previous reports of activity on NF-κB. They used in-house software (DESMOL11) and a commercial software (Dragon) for calculation of topological descriptors, which were used to build linear discriminant functions (classification models). The obtained model included a topological charge index and several connectivity indices, and achieved an overall accuracy of 84% on the training set. While the classification accuracy decreased significantly for a small 26-compound test set, the false positive rate remained very low, suggesting a high specificity model.

Interestingly, the authors reported, in the same article, a second topological discriminant function derived from a training set of 15 compounds active at UC (with no regard to their mechanism of action) and 16 presumably inactive compounds. In other words, *a phenotypic response and not a definite molecular event was used to define the dependent variable of the discriminant function*. The model calibrated with such training set achieved more than 80% accuracy on a 22-compound external test set. It included four topological indices: the eigenvalue sum from electronegativity weighted distance matrix, a Geary autocorrelation weighted by atomic polarizabilities, a connectivity index and a distance/detour index encoding information on fuse rings, and molecules' cyclicity. Both the NF-κB and the mechanism-independent model were later applied in a VS campaign of the Merck index database and the Microsource Pure Natural Product Collections.

The authors have used this same strategy (implementing UC drug discovery campaigns that combine the simultaneous application of mechanism-oriented models with mechanism-independent models) in other studies, as the reader will observe later in the text. Obtaining a model from a set of active compounds that do not share a mechanism of action is

a bold and controversial decision that deserves some discussion.

Traditionally, the idea that all the training examples used to infer a QSAR model should share the same mechanism of action (and the same binding mode) seemed rooted within the QSAR community. It appeared to be a reasonable condition in the light of certain three-dimensional (3D) QSAR methods meant to establish, in an indirect way, the necessary features for the ligand-target recognition event to occur [61–64]. It was argued that 3D QSAR methods were conceived to describe only one interaction step in the lifetime of ligands [61], a statement supported by the fact that many 3D QSAR methods are highly alignment-dependent. QSAR models would try to encode molecular features favorable and non-favorable to the pursued activity on the basis of some structural commonality (not necessarily an obvious one) across the calibration examples. After all, at the core of QSAR theory lies pattern recognition. The shared features, presumably, arose from the fact that all the considered ligands could bind the same target.

What is more, it was believed that only *in vitro* biological data should be considered, since *in vivo* data (and for that matter, phenotypic responses) reflect parallel processes (e.g. transport, metabolism, binding to multiple targets) and by definition is not possible to reach equilibrium in an *in vivo* system [61,62]. *In vitro* data is 'cleaner' than *in vivo* data, in the sense that interpretation of an *in vitro* assay is more straightforward and less affected by confounding factors or time-dependent changes. However, reductionist approaches could be dangerous when dealing with complex disorders (for instance UC).

There are many good reasons to take orthodox QSAR practices with caution.

First, very frequently biological data emerging from phenotypic models (e.g. *in vivo* or cellular models) are used to obtain QSAR models and even so the models achieve considerable explanatory and predictive power (see, for instance [65–75], or the study of Galvez-Llompart et al. under discussion).

Second, it is now understood that complex disorders (e.g. mood disorders, neurodegenerative conditions, epilepsy) are usually better addressed with drugs with complex pharmacology [76–78]. Some authors believe that the 'one target, one drug' paradigm has been disappointing in terms of innovative treatments [79–81]. Note that many of the conventional and last-generation treatments (biologics) for UC, including investigational drugs, are in agreement with the notion that some health conditions should be approached from a holistic perspective to reestablish a healthy state. From corticosteroids to monoclonal antibodies directly or indirectly targeting transcription factors (e.g. NF-κB), effective drugs in UC trigger a cascade of complex events that regulate several inflammatory mediators simultaneously. Some recent examples of applications of the systems pharmacology philosophy in the field of UC can be quoted. Muraro and Simmons developed an evolutionary optimization algorithm integrating known molecular interactions with gene expression data [82], which enable identification of differentially regulated network modules. The algorithm was applied to study experimental data derived from microarray analysis of Crohn's disease and UC biopsies and human interactome databases. The analysis allowed the extraction of dys-regulated subnetworks under different experimental conditions (inflamed and non-inflamed tissues in both diseases). The selected subnetworks included genes and pathways of known relevance for inflammatory bowel disease and revealed cross-talk among enriched pathways, mainly the Janus kinase/signal transducer (JAK/STAT) signaling pathway and the epidermal growth factor (EGF) receptor signaling pathway. Furthermore, integration of gene expression with molecular interaction data highlighted nodes that, although not being differentially expressed, interact with differentially expressed nodes and are part of pathways relevant to inflammatory bowel disease. Emig et al. have used a network-based approximation to propose new drug targets for 30 conditions, among them Crohn's disease and UC [83]; they demonstrated that their approach could possibly be used to detect new drug repurposing opportunities. Under this perspective, going against the doctrine and using phenotypic data might be in fact a better approach toward VS for novel agents against UC.

Third, most topological indices are conformation- and alignment-independent, and they are capable of describing more general properties than those relevant for single binding event. A substantial body of literature reporting successful developing of mechanism-independent topological models (using active training instances classified according to phenotypic responses) to identify drugs from a wide diversity of therapeutic categories (antivirals, anticonvulsants, bronchodilators, cytostatic drugs, antibacterials, antifungals, hypoglycemic, to mention just a few examples), suggest that, no matter the mechanism of action, drugs belonging to a particular therapeutic category tend to share common topological features [84–93].

Finally, QSAR theory has greatly evolved in the last years; multitasking QSAR models are suitable to predict multiple features, exploiting latent commonalities across tasks [94]. All in all, there seems to be no good reason to exclude a model where the mechanism is not known or when there are multiple mechanisms operating [95], especially taking into consideration the cumulative empirical evidence on MT ability to predict therapeutic class without mechanistic insight.

4.2. Interleukins 6 and 8

Matsumoto et al. have described the role of the inflammatory cytokine IL-6 in the establishment of chronic colitis [96] and elevated levels of IL-6 have been observed in patients suffering from inflammatory bowel disease [97]. The small molecule TAK-114, which downregulates TNF-α, IL-1, and IL-6, has reached clinical trials as potential treatment for UC. The humanized monoclonal antibody tocilizumab, which targets the IL-6 receptor, has been proposed as a potential new treatment for Crohn's disease [98]; recently, an anecdotical case of a patient with UC and rheumatoid arthritis that received tocilizumab and showed improvement was reported [99], although exacerbation after salvage therapy has also been reported [100].

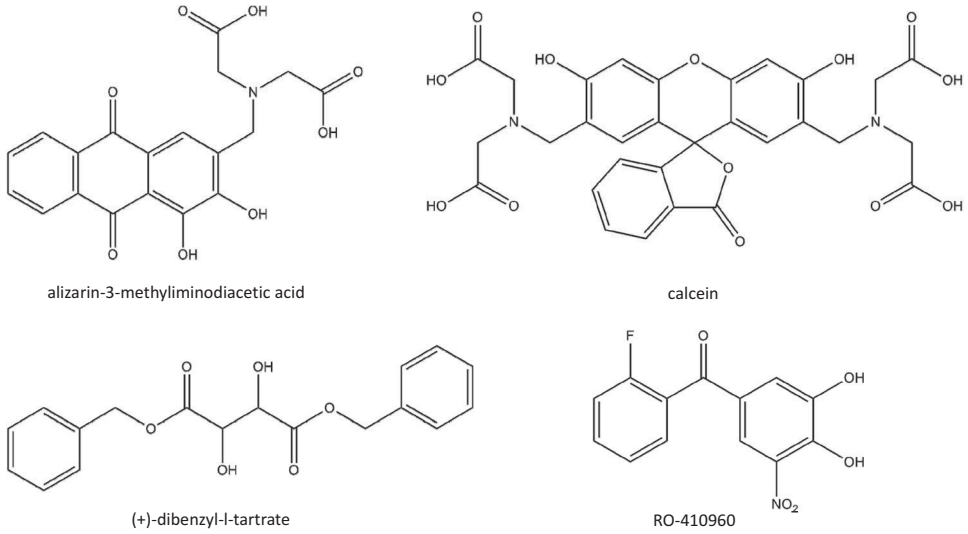


Figure 4. Potential UC treatments selected by molecular topology.

A similar approach to the one described in Section 4.1 has been applied to find novel inhibitors of IL-6 [29].

The topological QSAR model was developed from 53 compounds (25 actives and 28 inactives). It included only 3 descriptors but achieved an impressive 96% accuracy on the training set, and around an 80% of accuracy on a 108-compound test set. The selected independent variables were the spectral moment 15 from edge adjacency matrix weighted by resonance integrals, and the eigenvalue sums from Van der Waals and electronegativity weighted distance matrixes. The model was applied in a small VS campaigns on 28 compounds with hypothetical activity against UC. Four of the compounds (alizarin-3-methyliminodiacetic acid, calcein, (+)-dibenzyl-l-tartrate, and Ro 41–0960) (Figure 4) were selected by the model and submitted to cellular models (Caco-2 cells and RAW 264.7 macrophages) where three of them inhibited IL-6 production (the most potent, in the high nanomolar to low micromolar range).

Other QSAR models related to different ILs, including topological descriptors but not specifically focused on UC treatments, have been reported. For example, Pourbasheer et al. reported a QSAR model for the prediction of IL-1 receptor-associated kinase 4 (IRAK-4) [101]. IRAKs are a family of serine/threonine kinases involved in cellular signaling downstream of IL-1, IL-18, and Toll-like receptors [102]. They are critical for the activation of signaling cascades such as NF-κB. The reported model was built and validated from a reasonably diverse 65-compound data set, which was split into a 53-compound training set and a 12-compound test set. CODESSA and Dragon descriptors were used as independent variables; a combination of genetic algorithms and multi-linear regression was applied to infer the model. The model included some topological descriptors, such as Randić shape index path/walk 5 and molecular walk count of order 9, in combination with other non-topological features. Whereas the model displayed good performance in both training and test sets, no practical applications of it have been reported. Similarly, Assadolahi et al. used a combination of genetic algorithms and partial least squares to predict inhibition of IL-8 receptor, beta [103]. One hundred and thirty structurally heterogenous antagonist of such receptor were compiled from literature, and this dataset was representatively split in a 108-compound training set and a 22-compound test set using the Kennard–Stones algorithm. It showed good performance at the internal and external validations, and clearly outperformed linear models (stepwise multilinear regression and partial least squares) without the genetic algorithm-feature selection step. The latent variables of the model included several topological descriptors, among them a number of connectivity indexes, Galvez topological charge indices, and topological autocorrelations. Molecular modifications were then introduced in the data set compounds and the best model was applied in the prediction of the IL-8 receptor inhibitory activity of the newly designed compounds, though no experimental confirmation was performed. A similar, simpler QSAR modeling study focused on IL-8 receptor antagonists had previously been reported by the same group [104] and, once again, the best model included a topological autocorrelation (Geary autocorrelation-lag 5 weighted by atomic Van der Waals volume) in combination with non-topological molecular descriptors.

4.3. Tumor necrosis factor alpha

TNF-α is a validated target in UC; as described in Section 1, infliximab, adalimunab, and golimunab are three approved biologics for the treatment of UC that target TNF-α.

Very recently, the development of a discriminant function to identify TNF-α inhibitors was reported [30]. Using a 189-compound training set including 27 active compounds and 162 inactive ones and a pool of Dragon and DESMOL descriptors, the authors obtained a 5-descriptor classifier exclusively based on topological indices. Four of these descriptors were topological autocorrelations (Broto–Moreau, Moran, and Geary autocorrelations), which seem to confirm the ability of this type of descriptors to correlate with different biological activities (note that the same type of descriptors has once and

again appeared in other previously reviewed models). The reported discriminant function was validated with a 53-compound test set. The classificatory accuracy in both the training and test sets showed that the model has better performance on the inactive than on the active category, suggesting better specificity than sensitivity. In the same article, the authors reported discriminant and multi-linear regression models for inducible nitric oxide synthase (iNOS) inhibitory activity and, interestingly, a discriminant function capable of identifying compounds with activity in the dextran sulfate sodium (DSS)-induced colitis model, that is, a phenotypic response (in line with the previous discussion in Section 4.1). Also interesting is the fact that the authors provided *in vitro* and *in vivo* validation to their models' predictions, a step that is often omitted in studies from other groups that have been reviewed here. First, they conducted a small VS campaign on the four drug candidates previously selected and tested *in vitro* as suppressors of IL-6 production (Figure 4) [29], as previously stated in Section 4.2. All four candidates were predicted as active by all the models (TNF-α, iNOS and DSS-induced colitis). Three of the compounds inhibited NO production in a concentration-dependent manner. The four compounds inhibited TNF-α production in RAW264.7 cells (three of them in a concentration-dependent manner). Regarding the DSS-induced colitis model, three of the candidates were able to prevent colon shortening and had a positive impact on the Disease Activity Index. One of them also prevented DSS-induced weight loss.

Remarkably, the authors, through their successive and interconnected studies, are studying multiple cooperative mechanisms of action of the same four compounds plus their effects in a rodent model of colitis, which is in line with the previous discussion on the need of an integrative, holistic therapeutic approach to drug discovery for complex health conditions (e.g. inflammatory bowel disease). Complex diseases probably require simultaneous interventions at multiple points for effective treatment [80].

Some other QSAR models related to TNF-α inhibition and including topological descriptors have been reported, though they were obtained from a much less chemically diverse training set. Jain and Agrawal reported three linear models entirely based on topological descriptors and functional group counts [105]. The model that best performed at the validation step included two topological autocorrelations (Moran and Geary autocorrelations). The explanatory power of the obtained models was rather modest if we take into account that they were inferred from a training set consisting in only 22 compounds. Zhang et al. reported two QSAR models to explain the effects of 32 curcumin analogs with anti-inflammatory activity on TNF-α and IL-6 [106]; the performance of both models was modest (r^2 = 0.73 and 0.81, respectively) but, remarkably, the TNF-α model included two topological autocorrelations (Broto–Moreau and Geary autocorrelations) and the authors concluded that 'the inhibition rate is highly related to the skeleton structure' (i.e. the MT).

4.4. Other targets

α4β7 integrins expressed on lymphocytes adhere to ligands on the endothelial surface, allowing for migration of lymphocytes to inflammation sites. Antiintegrin antibodies (e.g. vedolizumab) inhibit the migration of lymphocytes across the mucosal barrier by targeting the α4β7 integrin [10]. Back in 2012, Jalali-Heravi and Mani-Varnosfaderani reported two Bayesian regularized genetic neural network to model the inhibition activity of 141 biphenylalanine derivatives as α4β7 and α4β1 integrin antagonists [107]. Both models achieved good performance on the training and test sets. Among several 3D descriptors and group counts, the α4β7 model included a topological autocorrelation. In the case of α4β1 inhibitory activity, the Randic shape index, the lowest eigenvalue of the Burden matrix, and the number of rotatable bonds were the most relevant model parameters. Soon later, Pourbasheer et al. reported linear and non-linear models of α1β4 inhibitory activity incorporating Dragon descriptors [108]. They used 41 moderately diverse compounds as training instances. The best results were observed using genetic algorithms as feature selection approach in combination with support vector machine, and the best model combined a topological descriptor (mean square distance index) among molecular descriptors from other categories.

Finally, QSAR modeling studies using MT have also been performed to search for mTOR inhibitors. mTOR is a serine-threonine kinase that regulates protein synthesis, cell growth, and proliferation in response to nutrients and growth factors; it is well established that mTOR plays a crucial role in tumorigenesis and accumulating evidence causally links increased mTOR activity to inflammatory responses [109]. mTOR inhibitors have shown positive effects in mice models of colitis [109,110]. Lakhlili et al. applied partial least squares to infer linear models from a data set of 364 molecules with inhibitory activity against mTOR in competition with ATP, extracted from PubChem Compound [111]. The data set was split into training (70%) and test (30%) sets. The model showed adequate performance on the training and test sets, and included topological descriptors such as Zagreb or Balaban indices. An *in silico* screening campaign was performed on a subset of around 2K compounds from ZINC database and a 1K-compound library from FDA fragment database. The study was complemented with docking simulations, but no experimental validation was performed.

5. Conclusion

To the moment, computer-aided drug discovery has provided very limited results in the field of UC, where drug discovery is centered on immunological studies. Whereas the number of computational studies addressing UC or validated drug targets in UC is quite scarce, most of the ligand-based approximations reported so far have resulted in QSAR models that include topological descriptors. Interestingly, some of the reviewed studies have included mechanism-blind models obtained from phenotypic data. A significant proportion of the reviewed reports include no experimental testing of drug candidates at all. In many cases, a narrow chemical diversity of the training sample is observed.

Various hits have emerged from *in silico* screening campaigns applying some of the reviewed models, and they have shown positive effects at the *in vitro* and *in vivo* level (DSS-induced colitis in mice).

6. Expert opinion

Considering the burden of the disease and the large number of ongoing clinical trials for potential new treatments (that reflect the interest in the discovery of novel therapeutics), relatively few applications of MT in combination with QSAR methods have been reported. In fact, the use of computer-aided approximations to identify new drug candidates for this disorder is very limited. What is more, a significant fraction of the reviewed studies does not include any kind of practical application or experimental validation of the reported models. This represents a general issue within the QSAR field. It should be noted that applications of QSAR models outside the research group in which they have been developed are extremely rare, with the exception of those included in commercial packages or freely available in public online resources. Having that in mind, it is important that QSAR modelers engage in practical applications of their models; otherwise, their work will likely remain strictly theoretic. The ultimate goal of the QSAR approach is to assist drug discovery campaigns, either explaining experimental observations or predicting them. Another caveat of some of the reviewed applications is that the reported QSAR models have been inferred from a training sample of limited size and/or chemical diversity, which severely limits the regions of the chemical space covered and the applicability of the resulting algorithms. As very often stated within the field, extrapolation should be avoided and predictions are reliable, at best, within the chemical space covered by the training data. A QSAR model derived from a single scaffold or chemotype will be reliably applicable to the activity prediction of drugs with such chemotype. Accordingly, the design of diverse training samples are crucial to obtain drug candidates with some degree of structural novelty.

The most consistent and comprehensive efforts in the field, in our opinion, come from the Molecular Connectivity and Drug Design Research Unit from the University of Valencia. Such statement is based in several observations. First, this research team has developed their models from molecular diverse training sets, which assures a wide applicability domain compatible with VS campaigns on large chemical libraries. Second, their models are exclusively based on topological descriptors, thus fully exploiting the advantages of topological descriptors (efficiency, conformation-independency, ease of calculation). Third, their contributions are the only ones among the reviewed articles that, to our knowledge, include some level of experimental validations of their models' predictions. Their work is not only theoretic and they have indeed contributed with novel lead compounds that have shown potential at the *in vitro* and *in vivo* level.

Some of the reviewed studies are in line with the network pharmacology philosophy, since the authors have combined *in silico* campaigns focused on different targets with models built on the basis of phenotypic responses. The integration of models using training data obtained from phenotypic, mechanism-independent experimental models of disease in combination with target-focused models is in agreement with the recent trend in the drug discovery arena to complement target-driven campaigns with classic phenotypic screening.

New active compounds have been found with multiple simultaneous effects (e.g. suppression of TNF-α and IL-6, inhibition of the NF-κB) and positive effects in the DSS-induced colitis model. It would be very interesting to see their efficacy assessed in other *in vivo* models of UC in the near future, in order to better judge their therapeutic potential. The multiple mechanisms of action of these candidates and their efficacy in the DSS-induced colitis models might not be casual. Due to the complexity of inflammation, known and investigational drugs for the treatment of UC target complex pathways, and even extremely selective therapies (e.g. monoclonal antibodies) exert their effects by targeting key nodes of inflammatory pathways that, indirectly, regulate the expression and activity of a diversity of biomolecules. Accordingly, we believe a rational approach to the discovery of novel therapeutics for UC should embrace the systems biology paradigm, either choosing highly selective agents that target key nodes of a metabolic pathway, or searching for multi-target agents that regulate multiple nodes simultaneously (e.g. bridging nodes). In agreement with this opinion, there are already some bioinformatics studies in the field of UC that have applied the networks perspective to identify dys-regulated network modules and potential new targets (e.g. hidden nodes responsible for maintenance of network connectivity). The future of the therapeutics against UC and any other complex disease probably lies in very selective treatments targeting highly connected nodes with the appropriate potency, or complex therapeutic solutions attacking multiple nodes, either as combined therapies or as multi-target agents (the later would perhaps be less prone to drug resistance issues). Microarrays technologies and bioinformatics will certainly be crucial to map and understand complex biological networks and restore health in multifactorial conditions.

An interesting (and, *a prior*, possibly counterintuitive) observation that emerges from a diversity of successful MT-based VS campaigns, including some of the efforts in the field of UC, is that MT seems to able to select new lead compounds for a specified therapeutic category without predefining a mechanism of action or narrowing the training samples of the model to those sharing a common mechanism of action. Such observation suggests that at least some therapeutic categories comprise drugs of similar topology (a therapeutic class topological pattern) independently of the mechanisms of action of each drug.

Finally, it can be highlighted that many of the reviewed models, no matter the pursued target, included topological autocorrelations as model features. This may reflect the importance of the molecular features to model UC-relevant effects, or more likely (taking into consideration the close relationship between this type of descriptors and the pharmacophore concept), the power of topological autocorrelations to capture relevant structural features for an efficient ligand-target interaction.

Funding

The authors are members of the Argentinian National Scientific and Technical Research Council (CONICET) and are researchers at the National University of La Plata (UNLP). These institutions fund the authors work.

Declaration of interest

The authors have no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

References

Papers of special note have been highlighted as either of interest (•) or of considerable interest (••) to readers.

1. Ordás I, Eckmann L, Talamini M, et al. Ulcerative colitis. Lancet. 2012;380:1606–1619.
   - **• An open access review on epidemiology, pathophysiology (including environmental and genetic factors) and valuable treatment algorithms at the moment of publication.**
2. Ananthakrishnan AN. Environmental risk factors for inflammatory bowel disease. Gastroenterol Hepatol (N Y). 2013;9:367–374.
3. Molodecky NA, Kaplan GG. Environmental risk factors for inflammatory bowel disease. Gastroenterol Hepatol (N Y). 2010;6:339–346.
4. De Souza HSP. Etiopathogenesis of inflammatory bowel disease: today and tomorrow. Curr Opin Gastroenterol. 2017;33:222–229.
   - **•• A comprehensive review on the etiology of inflammatory bowel disease.**
5. Cleynen I, Boucher G, Jostins L, et al. Inherited determinants of Crohn's disease and ulcerative colitis phenotypes: a genetic association study. Lancet. 2016;387:156–167.
   - **• A large scale (in fact, the larger so far) genotype-phenotype study on Crohn's disease and ulcerative colitis.**
6. Danese S, Fiocchi C. Ulcerative colitis. N Engl J Med. 2011;365:1713–1725.
7. Hindryckx P, Jairath V, D' Haens G. Acute severe ulcerative colitis: from pathophysiology to clinical management. Nat Rev Gastroenterol Hepatol. 2016;13:654–664.
8. Fuss I, Strober W. Ulcerative colitis. In: Mucosal immunology. 4th ed.Boston (MA): Academic Press; 2015.
9. Pastorelli L, Pizarro TT, Cominelli F, et al. Emerging drugs for the treatment of ulcerative colitis. Expert Opin Emerging Drugs. 2009;14:505–521.
10. Dubinsky MC. Reviewing treatments and outcomes in the evolving landscape of ulcerative colitis. Postgrad Med. 2017;129:538–553.
    - **•• A balanced and exhaustive review on the current treatments of ulcerative colitis and the outcome of clinical trials related to them.**
11. Caprilli R, Cesarini M, Angelucci E, et al. The long journey of salicylates in ulcerative colitis: the past and the future. J Crohns Colitis. 2009;3:149–156.
12. Rosseaux C, Lefebvre L, Dubuquoy L, et al. Intestinal anti-inflammatory effect of 5-aminosalicylic acid is dependent on peroxisome proliferator-activated receptor-gamma. J Exp Med. 2005;201:1205–1215.
13. Dubuquoy L, Rosseaux C, Thueru X, et al. PPARγ as a new therapeutic target in inflammatory bowel diseases. Gut. 2006;55:1341–1349.
14. Jannin V, Lemagnen G, Gueroult P, et al. Rectal route in the 21st Century to treat children. Adv Drug Deliv Rev. 2014;73:34–49.
15. Kornbluth A, Sachar DB; Practice Parameters Committee of the American College of Gastroenterology. Ulcerative colitis practice guidelines in adults: American College of Gastroenterology, Practice Parameters Committee. Am J Gastroenterol. 2010;105:501–523.
16. Faubion WA Jr, Loftus EV Jr, Harmsen WS, et al. The natural history of corticosteroid therapy for inflammatory bowel disease: a population-based study. Gastroenterology. 2001;121:255–260.
17. Toruner M, Loftus EV Jr, Harmsen WS, et al. Risk factors for opportunistic infections in patients with inflammatory bowel disease. Gastroenterology. 2008;134:929–936.
18. Park SC, Jeen YT. Current and emerging biologics for ulcerative colitis. Gut Liver. 2015;9:18–27.
19. Petkau JMV, Eksteen B. Selective biologics for ulcerative colitis and Crohn's disease - clinical utility of vedolizumab. Biologics. 2016;10:33–52.
20. Stawowczyk E, Kawalec P. Cost-effectiveness of biological treatment of ulcerative colitis – a systematic review. Prz Gastroenterol. 2017;12:90–97.
21. Khanna R, Jairath V, Vande Casteele N, et al. Efficient early drug development for ulcerative colitis. Gastroenterology. 2016;150:1056–1060.
22. Gálvez J, Gálvez Llompart M, García-Domenech R. Molecular topology as a novel approach for drug discovery. Expert Opin Drug Discov. 2012;7:133–153.
23. Amigó JM, Gálvez J, Villar VM. A review on molecular topology: applying graph theory to drug discovery and design. Naturwissenschaften. 2009;96:749–761.
24. Gálvez J, García-Domenech R. On the contribution of molecular topology to drug design and discovery. Curr Comput Aided Drug Des. 2010;6:252–268.
25. Wang T, Wu M-B, Lin J-P, et al. Quantitative structure–activity relationship: promising advances in drug discovery platforms. Expert Opin Drug Discov. 2015;10:1283–1300.
26. Cherkasov A, Muratov EN, Fourches D, et al. QSAR modeling: where have you been? Where are you going to? J Med Chem. 2014;75:4977–5010.
27. García-Domenech R, Gálvez-Llompart M, Zanni R, et al. QSAR methods for the discovery of new inflammatory bowel disease drugs. Expert Opin Drug Discov. 2013;8:933–949.
28. Gálvez-Llompart M, Recio MC, García-Domenech R. Topological virtual screening: A way to find new compounds active in ulcerative colitis by inhibiting NF-κB. Mol Divers. 2011;15:917–926.
29. Gálvez-Llompart M, Recio Iglesias MC, Gálvez J, et al. Novel potential agents for ulcerative colitis by molecular topology: suppression of IL-6 production in Caco-2 and RAW264.7 cell lines. Mol Divers. 2013;17:573–593.
30. Gálvez-Llompart M, Recio MC, García-Domenech R, et al. Molecular topology: A strategy to identify novel compounds against ulcerative colitis. Mol Divers. 2017;21:219–234.
31. Mayr LM, Bojanic D. Novel trends in high-throughput screening. Curr Opin Pharmacol. 2009;9:580–588.
32. Martins EA, Radhakrishan R, Badve RR. High-throughput screening: the hits and leads of drug discovery – an overview. J Appl Pharm Sci. 2011;1:2–10.
33. Fang Y, Ding Y, Feinstein WP, et al. GeuxDock: accelerating structure-based virtual screening with heterogeneous computing. PLoS One. 2016;11:e0158898.
34. Ge H, Wang Y, Li C, et al. Molecular dynamics-based virtual screening: accelerating the drug discovery process by high-performance computing. J Chem Inf Model. 2013;53:2757–2764.
35. Kristensen TG, Nielsen J, Pedersen CNS. Methods for similarity-based virtual screening. Comput Struct Biotechnol J. 2013;5:e201302009.
36. Cereto-Messagué A, Ojeda MJ, Valls C, et al. Molecular fingerprint similarity search in virtual screening. Methods. 2015;71:58–63.
37. Brown N, Lewis RA. Exploiting QSAR methods in lead optimization. Curr Opin Drug Discov Devel. 2006;9:419–424.
38. Miyao T, Kaneko H, Funatsu K. Inverse QSPR/QSAR analysis for chemical structure generation (from y to x). J Chem Inf Model. 2016;56:286–299.
39. McGaughey G, Sheridan R, Bayly C, et al. Comparison of topological, shape, and docking methods in virtual screening. J Chem Inf Model. 2007;47(4):1504–1519.
40. Talevi A, Bellera C, Castro E, et al. A successful virtual screening application: prediction of anticonvulsant activity in MES test of widely used pharmaceutical and food preservatives methylparaben and propylparaben. J Comput Aid Mol Des. 2007;21(9):527–538.
41. Lara-Valderrábano L, Rocha L, Galván E. Propylparaben reduces the excitability of hippocampal neurons by blocking sodium channels. NeuroToxicol. 2016;57:183–193.

42. Lara-Valderrábano L, Galván E, Rocha L. Propylparaben suppresses epileptiform activity in hippocampal CA1 pyramidal cells in vitro. Epilepsy Res. 2017;136:126–129.
43. Santana-Gómez C, Orozco-Suárez S, Talevi A, et al. Propylparaben applied after pilocarpine-induced status epilepticus modifies hippocampal excitability and glutamate release in rats. NeuroToxicol. 2017;59:110–120.
44. Bellera C, Balcazar D, Vanrell M, et al. Computer-guided drug repurposing: identification of trypanocidal activity of clofazimine, benidipine and saquinavir. Eur J Med Chem. 2015;93:338–348.
45. Sbaraglini M, Bellera C, Fraccaroli L, et al. Novel cruzipain inhibitors for the chemotherapy of chronic Chagas disease. Int J Antimicrob Agents. 2016;48(1):91–95.
46. Chen H, Yang Y, Engkvist O. Molecular topology analysis of the differences between drugs, clinical candidate compounds, and bioactive molecules. J Chem Inf Model. 2010;50(12):2141–2150.
47. Contrera J, Matthews E, Kruhlak N, et al. Estimating the safe starting dose in phase I clinical trials and no observed effect level based on QSAR modeling of the human maximum recommended daily dose. Regul Toxicol Pharmacol. 2004;40(3):185–206.
48. Zanni R, Galvez-Llompart M, García-Domenech R, et al. Latest advances in molecular topology applications for drug discovery. Expert Opin Drug Discov. 2015;10:945–957.
•• **A very interesting review that goes through the most recent advances and discussions regarding molecular topology.**
49. Randic M. Characterization of molecular branching. J Am Chem Soc. 1975;97:6609–6615.
50. Wiener H. Structural determination of paraffin boiling points. J Am Chem Soc. 1947;69:17–20.
51. Talevi A, Bruno-Blanch LE. Virtual screening applications in the search of novel antiepileptic drug candidates. In: Talevi A, Rocha L. Antiepileptic drug discovery. New York (NY): Springer; 2016.
52. Guimarães MC, Duarte MH, Silla JM. Is conformation a fundamental descriptor in QSAR? A case for halogenated anesthesics. Beilstein J Org Chem. 2016;12:760–768.
53. Furtula B, Lekishvili G, Gutman I. A graph theoretical approach to cis/trans isomerism. J Serbian Chemi Soc. 2014;79:805–813.
54. Todeschini R, Consonni V. Handbook of molecular descriptors. Weinheim: Wiley-VCH; 2008.
55. Hanser T, Barber C, Marchaland JF, et al. Applicability domain: towards a more formal definition. SAR QSAR Environ Res. 2016;27:893–909.
56. Ellison CM, Sherhod R, Cronin MT, et al. Assessment of methods to define the applicability domain of structural alert models. J Chem Inf Model. 2011;51:975–985.
57. Qin C, Zhang C, Xu F, et al. Theraeutic target database update 2014: A resource for targeted therapeutics. Nucleic Acids Res. 2014;42: D1118–D1123.
58. Sandborn WJ, Su C, Sands BE. Tolfacitinib as induction and maintenance therapy for ulcerative colitis. N Engl J Med. 2017;376:1723–1736.
59. Danese S. New therapies for inflammatory bowel disease: from the bench to the bedside. Gut. 2012;61(6):918–932.
60. Cătană C, Neagoe I, Cozma V, et al. Contribution of the IL-17/IL-23 axis to the pathogenesis of inflammatory bowel disease. World J Gastroenterol. 2015;21(19):5823–5830.
61. Teng M, Bowman E, McElwee J, et al. IL-12 and IL-23 cytokines: from discovery to targeted therapies for immune-mediated inflammatory diseases. Nat Med. 2015;21(7):719–729.
62. Oeckinghaus A, Ghosh S. The NF-κB family of transcription factors and its regulation. Cold Spring Harb Perspect Biol. 2009;1:a000034.
63. Lawrence T. The nuclear factor NF-κB pathway in inflammation. Cold Spring Harb Perspect Biol. 2009;1:a001651.
64. Langer T, Byrant SD. 3D quantitative structure-property relationships. In: Wermuth C. The practice of medicinal chemistry. London: Academic Press; 2008.
65. Sippl W. 3D QSAR: applications, recent advances, and limitations. In: Puzyn T, Leszczynski J, Cronin MT. Recent advances in QSAR studies. Method and applications. Dordrecht: Springer; 2010.
66. Roy K, Kar S, Das RN. Understanding the basics of QSAR for applications in pharmaceutical sciences and risk assessment. London (UK): Academic Press; 2015. Chapter 1, Background of QSAR and historical developments, p. 1-46.
67. Alagarsamy V. Textbook in medicinal chemistry. Vol. 1. New Delhi (India): Elsevier; 2013. Chapter 4, Structure-activity relationship and quantitative structure-activity relationship, p. 56-65.
68. Nikolic K, Agbaba D. QSAR study and design of novel organoselenium and α-tocopherol derivatives with enhanced chemotherapeutic activity. Lett Drug Des Discov. 2009;6:228–235.
69. Bello-Ramirez AM, Buendía-Orozco J, Nava-Ocampo AA. A QSAR analysis to explain the analgesic properties of aconitum alkaloids. Fundam Clin Pharmacol. 2003;17:575–580.
70. Patel SR, Gangwal R, Sangamwar AT, et al. Synthesis, biological evaluation and 3D-QSAR study of hydrazide, semicarbazide and thiosemicarbazide derivatives of 4-(adamantan-1-yl)quinoline as anti-tuberculosis agents. Eur J Med Chem. 2014;85:255–267.
71. Richard AM, Benigni R. AI and SAR approaches for predicting chemical carcinogenicity: survey and status report. SAR QSAR Environ Res. 2002;13:1–19.
72. Myshkin E, Brennan R, Khasanova T, et al. Prediction of organ toxicity endpoints by QSAR modeling based on precise chemical-histopathology annotations. Chem Biol Drug Des. 2012;80:406–416.
73. Tasso SM, Moon S, Bruno-Blanch LE, et al. Characterization of the anticonvulsant profile of valpromide derivatives. Bioorg Med Chem. 2004;12:3857–3869.
74. Shen M, LeTiran A, Xiao Y, et al. Quantitative structure–activity relationship analysis of functionalized amino acid anticonvulsant agents using k nearest neighbor and simulated annealing PLS methods. J Med Chem. 2002;45:2811–2823.
75. Talevi A, Enrique AV, Bruno-Blanch LE. Anticonvulsant activity of artificial sweeteners: a structural link between sweet-taste receptor T1R3 and brain glutamate receptors. Bioorg Med Chem Lett. 2012;22:4072–4074.
76. Roth BL, Sheffler DJ, Kroeze WK. Magic shotguns versus magic bullets: selectively non-selective drugs for mood disorders and schizophrenia. Nat Rev Drug Discov. 2004;3:353–359.
77. Bianchi MT, Pathmanathan J, Cash SS. From ion channels to complex networks: magic bullet versus magic shotgun approaches to anticonvulsant pharmacotherapy. Med Hypotheses. 2009;72:297–305.
78. Talevi A. Multi-target pharmacology: possibilities and limitations of the "skeleton key approach" from a medicinal chemist perspective. Front Pharmacol. 2015;6:20564.
79. Sams-Dodd F. Target-based drug discovery: is something wrong? Drug Discov Today. 2005;10:139–147.
80. Hopkins AL. Network pharmacology: the next paradigm in drug discovery. Nat Chem Biol. 2008;4:682–690.
81. Swinney DC, Anthony J. How were new medicines discovered? Nat Rev Drug Discov. 2011;10:507–519.
82. Muraro D, Simmons A. An integrative analysis of gene expression and molecular interaction data to identify dys-regulated sub-networks in inflammatory bowel disease. BMC Bioinformatics. 2016;17(1):art. no. 42.
83. Emig D, Ivliev A, Pustovalova O, et al. Drug target prediction and repositioning using an integrated network-based approach. PLoS ONE. 2013;8(4):art. no. e60618.
84. De Julian Ortiz JV, Galvez J, Muñoz-Collado C, et al. Virtual combinatorial syntheses and computational screening of new potential anti-herpes compounds. J Med Chem. 1999;42:3308–3314.
85. Bruno-Blanch LE, Galvez J, García-Domenech R. Topological virtual screening: a way to find new anticonvulsant drugs from chemical diversity. Bioorg Med Chem Lett. 2003;13:2749–2754.
86. Rios-Santamarina I, García-Domenech R, Galvez J. New bronchodilators selected by molecular topology. Bioorg Med Chem Lett. 1998;8:477–482.
87. Galvez J, García-Domenech R, Gomez-Lechon M, et al. Use of molecular topology in the selection of new cytostatic drugs. J Mol Struct Theochem. 2000;504:241–248.
88. De Gregorio Alapont C, García-Domenech R, Galvez J, et al. Molecular topology: a useful tool for the search of new antibacterials. Bioorg Med Chem Lett. 2000;10:2033–2036.
89. García-Domenech R, Ríos-Santamarina I, Catalá A, et al. Application of molecular topology to the prediction of antifungal activity for a

set of dication-substituted carbazols, furans and benzimidazoles. J Mol Struct Theochem. 2003;624:97–107.
90. Calabuig C, Antón-Fos GM, Galvez J, et al. New hypoglycaemic agents selected by molecular topology. Int J Pharm. 2004;278:111–118.
91. Talevi A. Network pharmacology and epilepsy. In: Talevi A, Rocha L. Antiepileptic drug discovery. New York (NY): Springer; 2016.
92. Marrero-Ponce Y, Ivarreta-Veitía M, Montero-Torres A, et al. Ligand-based virtual screening and in silico design of new antimalarial compounds using nonstochastic and stochastic total and atom-type quadratic maps. J Chem Inf Model. 2005;45:1082–1100.
93. Gavernet L, Talevi A, Castro E, et al. A combined virtual screening 2D and 3D QSAR methodology for the selection of new anticonvulsant candidates from a natural product library. QSAR Comb Sci. 2008;27(9):1120–1129.
94. Speck-Panche A, Cordeiro MN. Multitasking models for quantitative structure-biological effect relationships: current status and future perspectives to speed up drug discovery. Expert Opin Drug Discov. 2015;10:245–256.
95. Cronin MTD. Quantitative structure-activity relationships (QSARs) – applications and methodology. In: Puzyn T, Leszczynski J, Cronin MT. Recent advances in QSAR studies. Method and applications. Dordrecht: Springer; 2010.
96. Matsumoto S, Hara T, Hori T, et al. Probiotic Lactobacillus induced improvement in murine chronic inflammatory bowel disease is associated with the down-regulation of pro-inflammatory cytokines in lamina propia mononuclear cells. Clin Exp Immunol. 2005;140:417–426.
97. Waldner MJ, Neurath MF. Novel cytokine-targeted therapies and intestinal inflammation. Curr Opin Pharmacol. 2009;9:702–707.
98. Venkiteshwaran A. Tocilizumab. MAbs. 2009;1:432–438.
99. Szeto MCH, Yalcin MD, Khan A, et al. Successful use of tocilizumab in a patient with coexisting rheumatoid arthritis and ulcerative colitis. Case Report Immunol. 2016;2016:7562123.
100. Atreya R, Billmeier U, Rath T, et al. First case reporto f exacerbated ulcerative colitis after anti-interleukin-6R salvage therapy. World J Gastroenterol. 2015;21:12963–12969.
101. Pourbasheer E, Riahi S, Ganiali MR, et al. Quantitative structure-activity relationship (QSAR) study of interleukin-1 receptor associated kinase 4 (IRAK-4) inhibitor activity by the genetic algorithm and multiple linear regression (GA-MLR) method. J Enzyme Inhib Med Chem. 2010;25:844–853.
102. Janssens S, Beyaert R. Functional diversity and regulation of different interleukin-1 receptor-associated kinase (IRAK) family members. Mol Cell. 2003;11:293–302.
103. Asadollahi T, Dadfarnia S, Shabani AMH, et al. QSAR models for CXCR2 receptor antagonists based on the genetic algorithm for data preprocessing prior to application of the PLS linear regression method and design of the new compounds using in silico virtual screening. Molecules. 2011;16:1928–1955.
104. Ghasemi JB, Zohrabi P, Khajehsharif H. Quantitative structure-activity relationship study of nonpeptide antagonists of CXCR2 using stepwise multiple linear regression analysis. Monatsch Chem. 2010;141:111–118.
105. Jain HK, Agrawal RK. Topological descriptors in modeling tumor necrosis factor alpha inhibitory activity of xanthines, pteridinediones and related compounds. Internet Electron J Mol Des. 2007;6:218–228.
106. Zhang Y, Zhao L, Wu J, et al. Synthesis and evaluation of a series of novel asymmetrical curcumin analogs for the treatment of inflammation. Molecules. 2014;19:7287–7307.
107. Jalali-Heravi M, Mani Varnosfaderani A. QSAR modelling of integrin antagonists using enhanced Bayesian regularised genetic neural networks. SAR QSAR Environ Res. 2011;22:293–314.
108. Pourbasheer E, Aalizadeh R, Ganjali MR, et al. QSAR study of α1β4 integrin inhibitors by GA-MLR and GA-SVM methods. Struct Chem. 2014;25:355–370.
109. Bhonde MR, Gupte RD, Dadarkar SD, et al. A novel mTOR inhibitor is efficacious in a murine model of colitis. Am J Physiol Gastrointest Liver Physiol. 2008;295:G1237–1245.
110. Hu S, Chen M, Wang Y. mTOR inhibition attenuates dextran sulfate sodium-induced colitis by suppressing T cell proliferation and balancing TH1/TmTOR inhibition attenuates dextran sulfate sodium-induced colitis by suppressing T cell proliferation and balancing TH1/TH17/Treg Profile. PLoS ONE. 2016;11:e0154564.
111. Lakhlili W, Yasri A, Ibrahimi A. Structure–activity relationships study of mTORkinase inhibition using QSAR and structure-based drug design approaches. Onco Targets Ther. 2016;9:7345–7353.