European Journal of Pharmaceutical Sciences 102 (2017) 188–195

Contents lists available at ScienceDirect

European Journal of Pharmaceutical Sciences

journal homepage: www.elsevier.com/locate/ejps

Surface area, volume and shape descriptors as a novel tool for polymer lead design and discovery

Jonatan Riis Christensen [a], Helena Meng-Lund [a], Holger Grohganz [a,*], Antti Poso [b], Tuomo Laitinen [b], Ossi Korhonen [b], Lene Jørgensen [a], Jari Pajander [a]

[a] University of Copenhagen, Faculty of Health and Medical Sciences, Department of Pharmacy, Universitetsparken 2, 2100 Copenhagen, Denmark
[b] University of Eastern Finland, School of Pharmacy, Yliopistonranta 1, P.O. Box 1627, FI-70211 Kuopio, Finland

ARTICLE INFO

Article history:
Received 2 December 2016
Received in revised form 1 March 2017
Accepted 10 March 2017
Available online 12 March 2017

Keywords:
Molecular descriptors
Multivariate analysis
Volsurf
In silico

ABSTRACT

In recent years, the demand and interest for functionalized polymers have increased for drug delivery purposes. Because of the increased interest, methods that can be used to predict physical and chemical properties of polymers prior to synthesis would be of high value for the design and development of novel polymer structures. Through use of molecular descriptors and Principal Component Analysis, this study explores the possibilities of using in silico methods for polymer design and characterization for property prediction. The results presented in this paper suggest that it is possible to produce a model, which can successfully distinguish between a set of both structurally similar and different polymers based on their surface properties.

1. Introduction

Over the last decades, advances in synthesis capacities, especially through combinatorial chemistry, have resulted in large molecular libraries of small molecule drug candidates. With the increase in the number of candidates available for synthesis increases the demand for compound screening. Methods which are able to run a large number of automated screening assays or computer-based predictions, are in high demand (Hajduk and Greer, 2007; van de Waterbeemd and Gifford, 2003; Walters et al., 1998). The ability to predict possible adverse reactions caused by drug candidates in very early stages of development is essential for pharmaceutical companies, as the discovery of adverse effects in late stage development can be costly. In this context, quantitative structure-activity relationship (QSAR) modelling has been used extensively. QSAR usually relies on molecular descriptors, which transform structural chemical information into quantitative data. Thus QSAR enables the user to both interpret information about a certain molecule, such as physical and chemical properties, as well as to predict desirable properties of new structures usually by comparison of these properties with those of a lead compound (Todeschini and Consonni, 2008a). This approach is used in many aspects, including drug design and drug discovery, where it allows researchers to relatively quickly screen a database of chemical structures for desirable properties, based on the structure of the compounds. Examples include anti-tumor studies (Fortuna et al., 2008; Fortuna et al., 2010; Nicolle et al., 2009) blood-brain-barrier permeation studies (Clark, 2003; Hakkarainen et al., 2012), pharmacokinetics (Berellini et al., 2009; Tintori et al., 2009), transporter affinity studies (Omkvist et al., 2010; Thorn et al., 2007) chemical drug-likeness studies (Clark and Pickett, 2000) and toxicity and safety assessments (Dearden, 2003). Multiple types of molecular descriptors have been developed, describing everything from relatively simple properties such as molecular charge distribution (Galvez et al., 1994), to highly complex quantum chemical information (Karelson et al., 1996). In general, descriptors can be classified as either 2D or 3D. The 2D descriptors contain some of the more simple descriptors such as the constitutional descriptors, which reflect only the chemical composition of the compound such as number of atoms, bonds and rings with no regard to geometry or electronic structure. The 3D descriptors are also able to describe the nature of the atomistic composition, but additionally contain information about the spatial properties of the molecule (Todeschini and Consonni, 2008b).

Originally, VolSurf descriptors are quantitative numerical descriptors that have been extracted from the information in molecular field maps (MIF) calculated on 3D structures of a molecule and it is obtained as follows: First, different MIFs are generated around the target molecule, using hydrophilic and hydrophobic probes. Subsequently,

* Corresponding author.
E-mail address: holger.grohganz@sund.ku.dk (H. Grohganz).

http://dx.doi.org/10.1016/j.ejps.2017.03.013
0928-0987/

descriptor values are calculated based on these 3D MIF spaces. The obtained descriptor values then describe the surface of a structure in defined properties such as shape and size, hydrophilicity and hydrophobicity, and hydrogen-bonding ability (Cruciani et al., 2000b; Mannhold et al., 2006). Such properties can be correlated with properties or bioactivities of known compounds, and used for the development of new compounds with similar features (Crivori et al., 2000; Cruciani et al., 2000a; Hoest et al., 2007).

In addition to understanding the properties of an active pharmaceutical ingredient, the choice of formulation also plays a crucial role in research and development of a modern drug product. Lately, the development of new formulations, which are actively targeted to specific sites, e.g. tumors or mucosal membranes, have received increased attention. Especially advances in micro- and nanotechnology have been a major subject of interest, due to promising results in cancer treatments, where particles have been actively targeted to tumor cells (Cho et al., 2008; Nie et al., 2007; Peer et al., 2007) to increase response. While these strategies have promising outlooks as treatments, they rely heavily not only on the API itself, but also on the delivery vehicle and thus require optimal excipient selection. Traditionally, little attention has been paid to functionalities of excipients and excipient development, but in the case of nanotechnology, emphasis has been put on polymeric design in particular. With the increased interest in new and more complex polymer compounds, the traditional way of experimental testing of each candidate appears inefficient. Therefore, the implementation of polymer property modelling could proof to have large benefits with regard to the design and development of novel polymer compounds. Insight into the properties of existing polymers as well as newly developed polymers is highly beneficial, since it allows for selection of optimal polymers or polymer blends, for a given task. Recently Principal Component Analysis (PCA) was applied to describe the polymeric precipitation inhibition of poorly water soluble drugs and correlate the findings with simple physicochemical descriptors of the polymers showing that polymer properties responsible for inhibition of drug precipitation could roughly be identified (Warren et al., 2013).

While the molecular descriptor approach has been implemented in a variety of situations, as described previously (Berellini et al., 2009; Clark, 2003; Clark and Pickett, 2000; Dearden, 2003; Fortuna et al., 2008; Fortuna et al., 2010; Hakkarainen et al., 2012; Nicolle et al., 2009; Omkvist et al., 2010; Thorn et al., 2007; Tintori et al., 2009), modelling of polymeric structures through molecular descriptors has yet to be achieved. From a modelling perspective, characterization and prediction of polymer properties and behavior has been deemed very difficult, due to polymers susceptibility to conformational changes depending on the surrounding media. Computational modelling usually relies on the use of structures that have been optimized with regard to conformation. This can potentially introduce a significant amount of conformation related variance, and make interpretation of the results difficult and even inaccurate. However if successful, modelling of polymeric properties can find use in a huge number of industries and have a large impact on development of new products. It could allow for a whole new approach to polymer design and discovery, combining in silico methods with material technology.

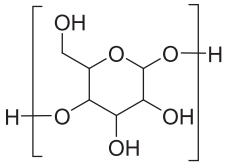
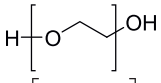
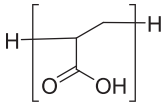
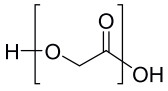
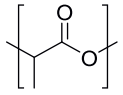
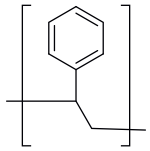
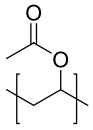
The scope of this paper is to investigate the potential of in silico methods combined with multivariate analysis by evaluating the applicability of VolSurf-like descriptors on a dataset of chemically different polymeric compounds that are commonly used and approved for use in humans. Furthermore, the most relevant molecular descriptors will be identified. The paper is based on the hypothesis, that through use of these volumetric surface descriptors, it will be possible to build a model that can successfully distinguish between the individual polymers based on the chemical structure, and that it will be able to do so, without special regard to conformation of the polymers. Ultimately, it aims to prove that it is possible to predict properties of the individual polymers using only a small fraction of the polymer chain, which enables the model to quickly and efficiently predict and characterize polymers for the design of novel structures.

2. Methods

2.1. Polymer simulations

A list of the nine polymers included in this study is shown in Table 1. Dassault Systèmes BIOVIA, Materials Studio, 5.0 (San Diego, US) was used to create structures for the polymers with chain lengths ranging from 1 up to 9 monomer units. Using the Condensed-phase Optimized Molecular Potentials for Atomistic Simulation Studies (COMPASS) force field in Materials studio, which is optimized for gas, liquid and condensed phase calculations, 101 different conformations with increasing energy for each of these structures were constructed using the Forcite module and Quench task. All simulations were done in constant temperature at 1000 K using the NVT ensemble with an integration time step of 1.00 fs, using the smart algorithm. The total simulation time was 100 ps over 100,000 steps. For every thousand steps, the conformation was stored and subjected to full geometry optimization. The structures were sorted according to potential energy, and from this database 10 conformations with low, medium and high-energy respectively, were chosen for further analysis. The first 10 structures being the ones with lowest potential energy were termed low energy

Table 1
Overview of the polymers included in the study, with abbreviations and structural formulas.

Polymer	Abbreviation	Structure
Maltodextrose	MD	
Poly (ethylene glycol)	PEG	
Poly carbonate	Carb	
Polyglycolic acid	PGA	
Polylactic acid	PLA	
Polystyrene	PSE	
Polyvinyl acetate	PVAc	
Polyvinyl alcohol	PVA	
Polyvinyl pyrrolidone	PVP