Journal of Controlled Release 173 (2014) 125–131

Contents lists available at ScienceDirect

Journal of Controlled Release

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

Computer-aided design of liposomal drugs: *In silico* prediction and experimental validation of drug candidates for liposomal remote loading

Ahuva Cern [a,b], Yechezkel Barenholz [a,*], Alexander Tropsha [c], Amiram Goldblum [b,**]

[a] Laboratory of Membrane and Liposome Research, Department of Biochemistry, IMRIC, The Hebrew University — Hadassah Medical School, Jerusalem, Israel
[b] Molecular Modeling and Drug Design Laboratory, The Institute for Drug Research, The Hebrew University of Jerusalem, Jerusalem, Israel
[c] The Laboratory for Molecular Modeling, UNC Eshelman School of Pharmacy, University of North Carolina, Chapel Hill, NC, USA

ARTICLE INFO

Article history:
Received 23 May 2013
Accepted 22 October 2013
Available online 31 October 2013

Keywords:
Liposomes
Remote loading
QSPR
Virtual screening
Iterative Stochastic Elimination
k-Nearest Neighbors

ABSTRACT

Previously we have developed and statistically validated Quantitative Structure Property Relationship (QSPR) models that correlate drugs' structural, physical and chemical properties as well as experimental conditions with the relative efficiency of remote loading of drugs into liposomes (Cern et al., J. Control. Release 160 (2012) 147-157). Herein, these models have been used to virtually screen a large drug database to identify novel candidate molecules for liposomal drug delivery. Computational hits were considered for experimental validation based on their predicted remote loading efficiency as well as additional considerations such as availability, recommended dose and relevance to the disease. Three compounds were selected for experimental testing which were confirmed to be correctly classified by our previously reported QSPR models developed with Iterative Stochastic Elimination (ISE) and k-Nearest Neighbors (kNN) approaches. In addition, 10 new molecules with known liposome remote loading efficiency that were not used by us in QSPR model development were identified in the published literature and employed as an additional model validation set. The external accuracy of the models was found to be as high as 82% or 92%, depending on the model. This study presents the first successful application of QSPR models for the computer-model-driven design of liposomal drugs.

1. Introduction

Liposomes present an effective drug delivery system (DDS): there are already more than 10 FDA-approved liposomal drugs [1,2] and many more preparations are in clinical trials [1]. A drug should fulfill several crucial conditions in order to become a successful liposomal product: (1) The concentration of the drug in the liposomes which is described as the drug to lipid (D/L) mole ratio should be high enough for administering a therapeutic dose. (2) The liposomal drug product should be stable at least at refrigerator temperature (4 °C). Product stability includes all parameters related to chemical and physical stability, which includes minimal drug leakage; stability at higher temperature will be an advantage. (3) Drug release in blood *in vivo* should occur slowly, enabling distribution of most of the liposomal drug to the target site.

For the purpose of parenteral administration, nano-size liposomes are mainly used [3,4]. The definition of nano-liposomes is based on structure and function. Regarding structure, it refers to liposomes in which there is a difference in the curvature between their two leaflets forming the lipid bilayer [3,4]. Regarding function, this is defined by biological considerations as only liposomes of size 100 nm or smaller benefit significantly from enhanced permeability and retention (EPR) and enhanced permeability (EP) effects and therefore are accumulating in cancerous and inflamed tissues [5]. In addition, the nano volume confers the liposomes with unique properties of highly efficient and stable drug loading as well as a controlled release profile. However, the paradox is that due to the very small internal volume of the nano liposomes they cannot be passively loaded with a therapeutically sufficient drug amount [6]. The approach of remote (active) loading was developed to overcome this obstacle and to achieve high drug concentrations in nano-liposomes [7–9]. This approach uses an ion gradient as the driving force for getting drugs into preformed liposomes to enable potentially high loading efficiency and good stability of the liposomal drug. Remote loading applies only to molecules that can accumulate in the internal aqueous phase of the liposome due to an ion or pH transmembrane gradient. Suitable candidates are amphipathic weak acids or weak bases having their logD (at pH 7) in the range of −2.5 to 2. Amphipathic weak bases should have a pKa ≤11 and weak acids should have pKa >3 [9]. Drug molecules that are too hydrophobic associate mainly with the lipid bilayer and will not be good candidates for remote loading [6]. On the other hand, molecules which are not amphipathic enough will not be remote loaded, as they will not be able to diffuse across the liposome lipid bilayer. Basic or acidic drug molecules suitable for remote loading can achieve equilibrium between the neutral, uncharged state, when a molecule can easily diffuse across the liposome membrane, and a charged

* Corresponding author. Tel.: +972 2 6757615; fax: +972 2 6757499.
** Corresponding author. Tel.: +972 2 6758701; fax: +972 2 6796584.
E-mail addresses: chezyb@ekmd.huji.ac.il (Y. Barenholz), amiram@vms.huji.ac.il (A. Goldblum).

0168-3659/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jconrel.2013.10.029

state, which in most cases prevents transport through the membrane. It is important to note that the success of this nanochemical engine is also due to the very small trapped aqueous volume of nano-liposomes, which supports faster and higher accumulation and intraliposome precipitation of a drug–counterion salt in crystalline or non-crystalline form. Amphipathic weak acids and bases can be effectively remote loaded to liposomes. However, the D/L ratios that will be loaded might be too low for administrating the therapeutic dose. For many drugs the therapeutic dose is relatively high (e.g., doxorubicin 50 mg/m^2, [10]) thus requiring a high D/L ratio in the formulation. In addition the formulation should maintain the D/L ratio during storage (i.e., minimal drug leakage) *In vivo* the release of drugs from the liposomes in the circulation should be slow, while it should be neither too slow nor too fast in the diseased tissue [6].

Liposomal formulation development requires considerable time and effort calling for the development of a computational modeling approach capable of predicting whether a drug is a good candidate for this DDS. To this end, we have begun recently to explore the utility of Quantitative Structure Property Relationship (QSPR) modeling as a computational tool to identify and prioritize drugs suitable for remote loading that satisfies the first and crucial condition for a good remote loading molecule, i.e., a high intra-liposomal drug concentration.

The first study to establish a correlation between drug structural properties and experimental conditions with remote liposome loading efficiency employed a decision tree method [9]; this model was built using data from the Barenholz laboratory. Additional data were generated recently for a larger set of drug molecules, and the enlarged data set was employed to develop novel QSPR models for *in silico* prediction of drug suitability for remote loading [11]. In formulating the objectives of the QSPR studies we took into account that the requirement of being amphipathic weak acids or bases is insufficient since some of such molecules could achieve high loading efficiency only at low D/L ratios, whereas other such compounds can be efficiently loaded at higher D/L ratios. This important observation suggests that other structural features, as well as experimental conditions, influence the D/L ratio that affords high loading efficiency. Thus, the feasibility for a drug to achieve high loading efficiency at relatively high initial D/L ratio used in the preparation was chosen as the target property to be predicted by QSPR models because this ratio determines whether a therapeutically effective dose could be administered using liposomal drug formulations. Taking into consideration that the amount of lipids allowed to be administered to a patient per day is limited, and that the D/L ratio determines the final drug concentration in the product, the latter determines the maximum dose that can be administered. Therefore, from a therapeutic standpoint we are ultimately interested in achieving high loading efficiency at relatively high initial D/L ratios. In our studies, the threshold of this initial D/L ratio was chosen to be 0.3, i.e., drugs that were capable of achieving high loading efficiency (greater than 70%) at D/L ratio of 0.3 or above were considered as "positive" candidates, whereas those that showed low loading efficiency above this threshold (even if they had high loading efficiency at D/L ratio below this threshold) were regarded as "negative" candidates.

The QSPR models to predict the initial D/L ratio were built using structural descriptors of drug molecules combined with experimental conditions, forming a joint set of compound characteristics as described in our previous publication [11]. Several machine learning techniques were employed to generate both the binary classification models (predicting "high" vs. "low" loading efficiency at D/L ratio $\geq$0.3) and continuous models (predicting the D/L ratio that enables high loading efficiency); these models were statistically validated using the 5-fold external validation technique discussed in detail previously [11].

In the present study we have used these previously developed QSPR models [11] for virtual screening of the Comprehensive Medicinal Chemistry (CMC) database containing in excess of 8000 molecules. Two of the earlier classification models were considered suitable for virtual screening for candidate compound selection; these models predicting D/L ratio from drug structural descriptors and experimental conditions were built using methods of Iterative Stochastic Elimination (ISE) and k-Nearest Neighbors (kNN), as discussed in our previous publication [11]. Three molecules emerging from virtual screening were prioritized for a proof-of-concept experimental validation in this study. Two of the selected drugs were predicted to be "positives" (i.e., predicted by both ISE and kNN models to have high loading efficiency at D/L $\geq$0.3) and one drug was predicted as "negative" (predicted by both models to have low loading efficiency at D/L $\geq$0.3). Each of the tested molecules was loaded to PEGylated liposomes by remote loading and analyzed for its loading efficiency at several D/L ratios using HPLC methods. The experiments showed that all three compounds were indeed correctly classified by the models. Furthermore, we have identified in the recently published literature [12–21] the remote loading data for 10 additional molecules that were not employed for model development in our previous study [11]. Nine of the drugs were reported as negatives and one was positive (using the same criteria as above). These 10 molecules were employed for additional validation of our models and the overall prediction accuracy for the external data set including both the three molecules tested by us and those reported in the literature was 92.3% and 81.8% for models built with ISE and kNN, respectively.

The vast amount of new data (from the literature of 2011–2012) emphasizes the importance of the remote loading method in the field of liposomal drug delivery. Additionally, finding only one drug with positive remote loading out of the 10 in recent literature emphasizes the difficulty of finding good candidates for remote loading into liposomes. Our studies confirm that QSPR models developed by our team are applicable for virtual database screening to identify candidates with high drug liposomal concentration achieved by the remote loading method. The models also predict correctly those molecules that are not good candidates for remote loading, i.e., the "true negatives". These observations underscore the power of the QSPR modeling method both to identify true positives and to reduce the number of false positives, saving experimental time, cost and efforts.

2. Materials and methods

2.1. Computational methods

2.1.1. Database to select candidate drugs for virtual screening

The CMC database includes more than 8000 molecules that are used or have been studied as medicinal agents in humans, pharmacological agents or biologically active compounds. This database was analyzed using the liposome remote loading models generated previously and described in Section 2.1.2. The models employed both structural/property descriptors (typical for any cheminformatics investigation) as well as experimental conditions (regarded effectively as additional experimental "descriptors" of compounds) for building models. The experimental conditions entered for all molecules included the liposome composition of hydrogenated soy phospholipid (HSPC):cholesterol:1,2-distearoyl-*sn*-glycero-3-phosphoethanolamine-*N*-[methoxy(polyethylene glycol)-2000] (mPEG DSPE) 55:40:5 mole ratio, loading duration of 10 min at 65 °C (above the T_m of the selected phospholipid) and 200 mM concentration of buffer used for gradient creation. Structural/property descriptors for each molecule were calculated using Molecular Operating Environment (MOE) software [22].

2.1.2. QSPR models

Two types of QSPR models to predict loading efficiency were built [11]: binary classification (category) and continuous models. Binary classification models were built using Iterative Stochastic Elimination (ISE), k-Nearest Neighbors (kNN) and Support Vector Machines (SVM). Continuous models were built using kNN and Support Vector Regression (SVR) (for additional details see ref [11]).

The binary classification models defined molecules as being "positive" when high remote loading efficiency ($\geq$70%) was achieved at D/L ratio $\geq$0.3. "Negatives" were those that showed low loading efficiency above