CORA Cork Open Research Archive
Cartlann Taighde Oscailte Chorcaí

Title	Applying computational predictions of biorelevant solubility ratio upon self-emulsifying lipid-based formulations dispersion to predict dose number
Authors	Bennett-Lenane, Harriet;Koehl, Niklas J.;O'Dwyer, Patrick J.;Box, Karl J.;O'Shea, Joseph P.;Griffin, Brendan T.
Publication date	2020-11-02
Original Citation	Bennett-Lenane, H., Koehl, N. J., O'Dwyer, P. J., Box, K. J., O'Shea, J. P. and Griffin, B. T. (2020) 'Applying computational predictions of biorelevant solubility ratio upon self-emulsifying lipid-based formulations dispersion to predict dose number', Journal of Pharmaceutical Sciences, 110(1), pp.164-175. https://doi.org/10.1016/j.xphs.2020.10.055
Type of publication	Article (peer-reviewed)
Link to publisher's version	10.1016/j.xphs.2020.10.055
Rights	© 2020, Elsevier B.V. All rights reserved. This manuscript version is made available under the CC BY-NC-ND 4.0 license. - https://creativecommons.org/licenses/by-nc-nd/4.0/
Download date	2026-10-13 13:23:07
Item downloaded from	https://hdl.handle.net/10468/10816

UCC
University College Cork, Ireland
Coláiste na hOllscoile Corcaigh

Applying Computational Predictions of Biorelevant Solubility Ratio upon Self-Emulsifying Lipid-Based Formulations Dispersion to Predict Dose Number.

Harriet Bennett-Lenane[1], Niklas J. Koehl[1], Patrick J. O'Dwyer[1,2], Karl J. Box[2] , Joseph P. O'Shea[1] , Brendan T. Griffin[1].

[1]School of Pharmacy, University College Cork, Cork, Ireland.

[2]Pion Inc. (UK) Ltd., Forest Row, East Sussex, UK.

*Correspondence to: Brendan Griffin; Tel.: +353 (0) 21 4901657; fax: +353 (0) 21 4901656. Email address: Brendan.Griffin@ucc.ie.

Abstract

Computational approaches are increasingly utilised in development of bio-enabling formulations, including self-emulsifying drug delivery systems (SEDDS), facilitating early indicators of success. This study investigated if *in silico* predictions of drug solubility gain i.e. solubility ratios (SR), after dispersion of a SEDDS in biorelevant media could be predicted from drug properties. Apparent solubility upon dispersion of two SEDDS in FaSSIF was measured for 30 structurally diverse poorly water soluble drugs. Increased drug solubility upon SEDDS dispersion was observed in all cases, with higher SRs observed for cationic and neutral versus anionic drugs at pH 6.5. Molecular descriptors and solid-state properties were used as inputs during partial least squares (PLS) modelling resulting in predictive models for SR_{MC} (r^2 = 0.81) and SR_{LC} (r^2 = 0.77). Multiple linear regression (MLR) facilitated generation of simplified SR equations with high predictivity (SR_{MC} r^2 = 0.74; SR_{LC} r^2 = 0.69), requiring only three drug properties; partition coefficient at pH 6.5 ($logD_{6.5}$), melting point (Tm) and aromatic bonds as fraction of total bonds (FArom_B). Through using the equations to inform drug developability classifications (DCS) for drugs that have already been licensed as lipid based formulations, merits for development with SEDDS was predicted for 2/3 drugs.

Abbreviations:

PWSD, Poorly Water-Soluble Drug; LBF, Lipid-Based Formulations; SEDDS, Self-emulsifying drug delivery systems; GIT, Gastrointestinal tract; O/W, oil-in-water; DCS, Developability Classification System; DCS, Developability Classification System; rDCS, Refined Developability Classification System; An Absorption Number; Do, Dose Number; BCS, Biopharmaceutics Classification System; $SEDDS_{Migylol812}$, Migylol 812 based SEDDS, $SEDDS_{OliveOil}$, Olive Oil based SEDDS.

Introduction

Increasing numbers of poorly water soluble drugs (PWSD) in development pipelines has intensified the need for bio-enabling formulations to enhance oral bioavailability (1, 2). One such approach involves administration of drug substances in Lipid-Based Formulations (LBFs), which enhance apparent solubility of PWSDs, while potentially also increasing absorption via stimulation of endogenous lipid absorption pathways for lipophilic xenobiotics. Despite numerous commercial examples of LBFs, with previous estimations of up to 4% of orally administered drug products utilising LBFs (3), it was recently observed that relative numbers of new commercial products using LBFs have declined over the last decade (4). Such statistics suggest challenges to more widespread adoption of LBFs among pharmaceutical companies, potentially linked to a lack of clear guidance on appropriate early screening to guide bio-enabling formulation selection (5).

Self-emulsifying drug delivery systems (SEDDS) fall under the umbrella term LBFs, and refer to combinations of oils with surfactants and co-solvents which spontaneously emulsify forming a stable emulsion on dispersion in the gastrointestinal tract (GIT) (6). Ability to self-emulsify and maintain solubilisation on dispersion is a key SEDDS performance determinant. Typically, the drug dose should be soluble in the SEDDS vehicle, and much effort is focused on determining the inherent lipid solubility of the drug, usually involving resource intensive drug solubility screening in a range of lipid excipients (7, 8). More recently, the application of computational models, in conjunction with drug biopharmaceutical profiling, has been explored to support higher throughput formulation selection for LBFs (9, 10). While certain computational approaches aim to determine drug properties which produce favourable oral drug candidates, other tools have instead examined molecular properties that may signal necessity for use of bio-enabling strategies or alternatively signal greater suitability for a particular type of bio-enabling strategy. Regarding the latter, a number of noteworthy studies have demonstrated utility of computational pharmaceutics approaches to predict lipid solubility to act as a guide for maximal dose loading in LBF pre-concentrates (11-14). Critically, while predictions of the drug solubility in lipids are useful to guide initial understanding of the maximum dose loading in the SEDDS vehicle, this approach does not represent the sole criterion for LBF suitability.

Modifications in the GIT upon SEDDS ingestion are crucial in determining formulation performance, as solubilisation capacity in luminal media can be altered dramatically following SEDDS emulsification and through interactions of lipid excipients with endogenous solubilising species (15-17). SEDDS dispersion leads to increased drug solubilisation, transient supersaturation, and potentially precipitation, thereby presenting drugs to intestinal fluids at concentrations exceeding their equilibrium solubility (18). From a biopharmaceutical perspective, apparent drug solubility in intestinal fluid upon SEDDS dispersion appears critical in determining LBF suitability. Accordingly, the lipid formulation-performance classification system emphasises formulation capability to retain solubilisation upon dispersion and digestion (19). The use of simulated biorelevant fluids in such *in vitro* assessments is likely to be a more reliable indicator of whether a SEDDS approach can effectively solubilise the dose *in vivo*. Biorelevant testing is an integral part of pharmaceutical characterisation, revealing concentrations likely to be soluble within human intestinal fluids (HIF) (20, 21), while a key tenet of the

Developability Classification System (DCS) is the use of biorelevant solubility in fasted state simulated intestinal fluids (FaSSIF) as an improved guide to *in vivo* performance and drug developability (22). More recently, a refined DCS (rDCS) extended this developability concept to include customised *in vitro* assessments of supersaturation and precipitation risks involving Absorption Number (An) and Dose Number (Do) (23). Such developability guides, along with decision trees utilising biorelevant media instead of buffered aqueous media (22, 24-26), signify emerging emphasis on developability and biopharmaceutical concepts in early product testing. However, as *in vitro* techniques utilised to predict the dose that is effectively solubilised *in vivo* can be complex and resource heavy, development of models capable of predicting this dose are strongly merited (27).

With regard to both advancing LBF computational pharmaceutics and use of biopharmaceutically relevant conditions, our hypothesis was to apply a computational approach to predict solubility increases upon SEDDS dispersion. Given the inherent complexity of the mixed colloidal species formed upon dispersion of SEDDS with endogenous biliary lipids, approaches to predict apparent solubility are considered complex at this stage. As an alternative, the solubility increase achieved via SEDDS dispersion in FaSSIF, relative to drug solubility in FaSSIF, represents a more realistic modelling parameter. This can be used to inform the maximal dose solubilised within the intestine, assuming experimental drug solubility in FaSSIF is known. Accordingly, this study attempted to apply a computational approach in relating drug properties to predict solubility increases (i.e. solubility ratios) following SEDDS dispersion. This approach can therefore be used to effectively guide the dose number (Do) produced in intestinal fluids. Subsequently, this study explored suitability of linking the predicted Do to the framework provided by the DCS and hence, providing a tool for guiding developability of a SEDDS formulation strategy in early stage drug development. To achieve this aim, apparent drug solubility of 30 PWSD was experimentally determined upon dispersion in FaSSIF of two prototype SEDDS. SEDDS were selected based on prior ability to spontaneously emulsifying, forming a stable microemulsion and were composed of either a medium chain ($SEDDS_{Migylol812}$) or long chain ($SEDDS_{OliveOil}$) oil phase, with a common surfactant, co-surfactant blend in order to examine their excipient effects (27). Solubility ratios (SR_{MC} and SR_{LC}) achieved versus FaSSIF solubility were collated with drug descriptors to develop computational models and predictive equations. Through prediction of DCS classifications, this work aimed to advance the concept of computational pharmaceutics to inform drug developability, exemplifying use of predictive tools to expedite formulation options in early development.

Methods

Dataset Selection

A dataset of 30 structurally diverse PWSD was selected (Table 1). Drugs were selected based on a range of criteria including availability of published reports of drug properties, utilisation in previous LBF computational modelling publications and drugs commercially licensed as LBFs (11, 12, 25, 28). Light absorbing ability of the compounds' UV-chromophores were also considered, to allow sufficient detectability by the fibre optic UV probes of the µDISS Profiler. A final drug data set was selected to

ensure a sufficient representation of drugs categorised as anionic (8), cationic (9) and neutral (13) overall at pH 6.5. The Henderson-Hasselbalch equation was used to determine ionisation at pH 6.5 (Table 1). All drug compounds were purchased from Kemprotec Ltd (Cumbria, United Kingdom). The final dataset displayed a wide physiochemical profile of molecular weight (MW) (230-868.44 g/mol), lipophilicity (clogP) (2.1-7.1) and melting point (T_m) (79-296.5ºC).

Formulations

Two SEDDS previously utilised for oral delivery of a model PWSD in preclinical studies were chosen (27). $SEDDS_{Migylol812}$ contained 40% w/w medium chain triglycerides (Miglyol 812) with 20% w/w surfactant (Kolliphor RH 40 - polyoxyl-40-hydrogenated castor oil) and 40% w/w co-surfactant (Tween 85 - polyoxyethylene-(20)–polysorbitan trioleate). $SEDDS_{OliveOil}$ contained 40% w/w long chain triglycerides (olive oil), while quantities of surfactant and co-surfactant remained similar to $SEDDS_{Migylol812}$, with 20% Kolliphor RH 40 and 40% Tween 85. Migylol 812N is primarily composed of C8 and C10 fatty acids (approx. 60:40%). Olive oil contains saturated and unsaturated fatty acids of primarily C16-C18 chain length. Miglyol 812N was kindly gifted from IOI Oleo GmbH (Hamburg, Germany), while Olive Oil, Tween 85 and Kolliphor RH 40 were purchased from Sigma-Aldrich (Ireland). SEDDS were prepared by weighing exact excipient quantities into a screw cap glass tube and incubated at 37 °C, overnight on a stirring plate 200 rpm (Mixdrive 15, 2MAG, Germany).

Media Preparation

Phosphate buffer ($PhB_{pH6.5}$) and FaSSIF-V1 were prepared according to biorelevant.com (Croydon, UK) protocol and adjusted to pH 6.5 using a Model 3510 pH/mV/Temperature Meter (Jenway, UK). FaSSIF-V1 was chosen due to high correlation with HIF and availability of drug solubility datasets (29, 30). Water was obtained from a MilliQ water system. All chemicals and solvents were of analytical or high-performance liquid chromatography (HPLC) grade and purchased from Sigma-Aldrich (Ireland). Conditions for simulating dispersion of the SEDDS in intestinal fluids were produced by dispersing SEDDS (1:200 dilution) in $PhB_{pH6.5}$ (i.e. $PhB_{pH6.5}$-$SEDDS_{Migylol812}$ and $PhB_{pH6.5}$-$SEDDS_{OliveOil}$) and FaSSIF (i.e. FaSSIF-$SEDDS_{Migylol812}$ and FaSSIF-$SEDDS_{OliveOil}$). This lipid dilution was chosen to be typical of reasonable lipid concentrations found in a biorelevant volume.

Media Characterisation: Media Droplet Size and Zeta Potential

Droplet size (nm) and polydispersity index (PDI) of FaSSIF, FaSSIF-$SEDDS_{Migylol812}$, FaSSIF-$SEDDS_{OliveOil}$, $PhB_{pH6.5}$-$SEDDS_{Migylol812}$ and $PhB_{pH6.5}$-$SEDDS_{OliveOil}$ were measured using Dynamic Light Scattering (DLS) with a Malvern Zetasizer Nano ZS (Malvern Analytical, US) with a 4mW 633nm He-He laser at 37°C with a backscattering angle of 173° using the Stokes-Einstein equation. Measurements were performed with unfiltered samples in disposable UV-cuvettes from Sarstedt AG & Co. KG (Numbrecht, Germany) (10 x 4 x 45 mm). Refractive indices used were 1.1333 ($PhB_{pH6.5}$) and 1.334 (FaSSIF) (31). The electrophoretic mobility i.e. ζ-potential, of colloidal structures in the media was measured using the Zetasizer in disposable folded capillary cells (DTS1070) using the Helmholtz-Smoluchowski equation (32). Each analysis was conducted in triplicate, presented as mean ± standard deviation.

Experimental Solubility Determination

Apparent drug solubility studies in FaSSIF, FaSSIF-SEDDS$_{Migylol812}$ and FaSSIF-SEDDS$_{OliveOil}$ were experimentally determined over 24 hours as the 24 hour time point was used for solubility ratios. Solubility was determined via either shake flask with RP-HPLC/UV analysis (6 drugs) or µDISS Profiler (Pion INC, Woburn, MA) (24 drugs), where preliminary studies verified method comparability (Supplementary Materials).

Shake Flask Method

Drug was added in excess to triplicate glass vials containing either FaSSIF, FaSSIF-SEDDS$_{Migylol812}$ or FaSSIF-SEDDS$_{OliveOil}$ (n=3). pH was maintained at 6.5 prior to experiments. Vials were placed on a stirring plate (Mixdrive 15, 2MAG, Germany) in a 37°C incubator at 300 rpm. 300 µl samples were removed at 2, 4, 6 and 24 hours. Excess solid was separated using a centrifuge for 15 minutes at 21,380 x g (Mikro 200 R, Andreas Hettich GmbH & Co. KG, Germany). Samples were diluted in acetonitrile for analysis via RP-HPLC/UV. Drug Detection was conducted using an Agilent 1200 series HPLC system. The columns and mobile phases used for each drug analysed along with injection volume, flowrate and detection wavelength can be found in supplementary materials.

µDISS Profiler

Apparent drug solubility (n = 3) was determined at a stirring rate of 300 rpm over 24 hours (37°C). Path length of the *in situ* UV probes was varied (1 - 5 mm) depending on anticipated concentration range and the UV absorbance properties of the drug molecule. Standard spectra were collected for each compound at pH 6.5 and a linear relationship ($r^2 > 0.99$) was established between absorbance and concentration in each case. The experimental run was performed in six vials where a large excess of API was added (10-20 times more than the anticipated FaSSIF solubility) to account for the potentially large solubility enhancement. These vials contained 15 ml FaSSIF-SEDDS$_{Migylol812}$ or FaSSIF-SEDDS$_{OliveOil}$ and a cross-bar magnetic stirrer. Two additional channels were used as blanks to consolidate for potential issues with background changing FaSSIF UV absorbance over time. The *in situ* UV probes scanned the samples at predefined time intervals (30 minutes). Concentrations were determined by considering area-under-the-curve (AUC) in second derivative spectra, to lessen interference from background turbidity. A range of wavelengths were utilised to quantify drug. Data was interpreted using the Au Pro software (Version 5, Pion INC, MA, USA).

Drug Physiochemical and Molecular Properties

In excess of 250 descriptors including physiochemical and modelling descriptors were obtained from ADMET Predictor 9.5 (Simulations Plus, USA). T_m was obtained from literature (11, 24, 28). Biorelevant solubility values in FaSSIF, FeSSIF and $PhB_{pH6.5}$ were obtained from literature sources where available (28, 33). In absence of published data, predicted solubility values were generated (ADMET Predictor, Ver. 9.5, Simulations Plus Inc., US). Highest licensed drug dosage strengths were obtained from the European Medicines Agency (EMA) or Food and Drug Administration (FDA) databases.

Biopharmaceutical Data Analysis

Apparent drug solubility values in all media are presented as mean ± standard deviation (n=3) (Supplementary Materials). Solubility ratios (SR) for the 30 drugs with either FaSSIF-SEDDS$_{Migylol812}$ (referred to as SR_{MC}) or FaSSIF-SEDDS$_{OliveOil}$ (referred to as SR_{LC}) versus FaSSIF were calculated via Equation 1:

$$(1)\ Solubility\ Ratio\ (SR) = \frac{Apparent\ Solubility\ FaSSIF-SEDDS}{Apparent\ Solubility\ FaSSIF}$$

SR standard error (SE) was calculated from Equation 2 as previously reported (24):

$$(2)\ \mathrm{SE} = SR\ x\ \sqrt{\frac{SA^2}{A^2} + \frac{SB^2}{B^2}}$$

Where A, B, SA and SB refer to the mean measured solubility values (24hrs) and standard errors for A (FaSSIF) and B (FaSSIF-SEDDS) respectively. In order to assess capacity for SEDDS to bridge the fasted-fed state solubility gap, SR_{MC} and SR_{LC} were related to comparative SRs for each drug using FeSSIF solubility in place of drug solubility upon SEDDS dispersion i.e. FeSSIF/FaSSIF. Graphs illustrating SRs were obtained using Prism (Version 5, Graphpad, USA). Linear regression was performed using Excel (Microsoft Office, 2016) to assess correlations between SR and individual drug properties or solubility in various media. To test significance between paired solubility values in FaSSIF-SEDDS$_{MC}$ versus FaSSIF-SEDDS$_{LC}$ the distribution of the difference was used to determine normality. A two sided bootstrap-paired test (5000 samples) was used to determine significance ($p < 0.05$). A simple scatter plot was produced for FaSSIF-SEDDS$_{Migylol812}$ versus FaSSIF-SEDDS$_{OliveOil}$ and regression coefficients fitted for interpretation and a bootstrap test for the coefficients conducted. A two sided independent samples t-test was used to analyse media droplet sizes and Levene's test was used to check for equality of variances. A p-value < 0.05 indicated a violation of equal variance. All statistical analysis was conducted using SPSS Statistics (Version 26, IBM Corporation, US).

Multivariate Data Analysis and Modelling Parameters

Multivariate data analysis (MDA) was conducted using Unscrambler (Version 11, Camo Analytics, US). Molecular structures were acquired as smiles from PubChem and used as inputs for the ADMET Predictor software (Version 9.5, Simulations Plus, California, USA) to calculate >250 molecular descriptors. These were added to PSA and T_m and used as variable inputs for Principal Component Analysis (PCA) and Partial Least Squares (PLS) modelling. Modelling responses were the logarithm of SR in both SEDDS ($\log SR_{MC}$ and $\log SR_{LC}$). PCA was first applied randomly to aid training/test set identification. A split of 70:30 (21:9 drugs) of training:test set was used to increase model robustness. Training set criteria was that it covered the test set chemical space along with a relatively even spread of SRs. Influential outliers were placed in the test set if they displayed both large residual and high leverage in the Influence plot. A Hotelling's T^2 ellipse was also applied for outlier detection (95% confidence interval).

PLS was used to establish important descriptors for predicting SR_{MC} and SR_{LC}. The nonlinear iterative partial least squares (NIPALs) algorithm was utilised and all 250+ variables were mean

centred, de-identified and standardized through scaling by standard deviation. Descriptors displaying the same value for all drugs were removed, along with skewed descriptors. To limit overfitting potential, a limit of two principal components was used. Variable reduction was performed to decrease complexity and noise. A Marten's uncertainty test was applied to help identify important variables and assess stability. This involved a "jackknifing" procedure and production of sub-models to identify non-significant variables (34). Variable weighted beta coefficient rankings from the Important Variables plot and their p-values were also used to remove unimportant variables. Variables in the same area of the correlation loadings plot were removed leaving a singular variable. Variables near the centre of this plot were removed. Any change in r^2 calibration and r^2 validation was monitored. Model accuracy was validated by the Root Mean Square Error (RMSE) of validation and calibration. Models were validated by a full cross validation (leave-one-out) to improve power and by test sets of drugs not used in model development to strengthen general applicability.

Solubility Equation for Predicting Biopharmaceutical Dose Number and DCS Class.

It was then investigated if easily interpretable equations based on drug properties could predict SR_{MC} and SR_{LC}. Multiple linear regression (MLR) was performed using Excel (Microsoft Office, 2016) to investigate correlations between selected significant PLS model variables versus $logSR_{MC}$ and $logSR_{LC}$. Equation development was monitored by descriptor p-values, the f-value, r^2 and adjusted r^2. The same training and test sets as PLS were used.

DCS classification of each drug was obtained using solubility and permeability parameters outlined previously (22, 23). While drug permeability was predicted from the ADMET Predictor (Version 9.5, Simulations Plus Inc., US), solubility criteria was obtained using a dose/solubility ratio in 500mls of media using equation (3):

$$(3)\ Do = \frac{Dose}{(Ssi)(Vsi)}$$

Where, Dose is the highest dose, S_{si} apparent solubility in biorelevant media i.e. FaSSIF, V_{si} is the available fluid volume for dissolution in the small intestine (500 ml).

Solubility criteria for DCS classification using experimentally determined solubility's upon SEDDS dispersion was calculated using equation (4):

$$(4)\ Do_{(SEDDS)} = \frac{Dose}{(Cs)(Vsi)}$$

Here, C_s is apparent drug solubility upon SEDDS dispersion in biorelevant media i.e. FaSSIF-SEDDS.

For DCS classifications using the predicted solubility ratios (SR) from the MLR equations, $Cs_{(Predicted)}$ was calculated using equation 5:

$$(5)\ Cs_{(Predicted)} = SR * Ssi$$

Where SR is the predicted solubility ratio upon SEDDS dispersion from the MLR equations and S_{si} is apparent solubility of the compound in biorelevant media i.e. FaSSIF. Incorporating equation 5, solubility criteria for DCS classifications upon SEDDS dispersion was predicted using equation 6:

$$(6)\ Do_{(Predicted)} = \frac{Dose}{(Cs_{(Predicted)})(Vsi)}$$

Predicting DCS Classifications of Commercial LBF Drugs

To assess the equations' general applicability to make predictions for drugs outside equation development and validation, $Do_{(Predicted)}$ was applied to a list of drugs that have been successfully commercialised as LBF products (4). DCS classifications were produced using $Do_{(predicted)}$ values (Equation 6), to predict if dose solubility limitations for the commercial drugs would be overcome upon SEDDS dispersion. FaSSIF solubility was obtained from literature or from the ADMET Predictor 9.5 (Simulations Plus, USA) (Fagerberg et al., Fagerberg and Bergstrom, 2015). Predicted classifications were compared to classifications using FaSSIF solubility alone (Equation 3).

Results

SEDDS Characterisation on Biorelevant Dispersion

$SEDDS_{Migylol812}$ and $SEDDS_{OliveOil}$ both dispersed in FaSSIF and $PhB_{pH6.5}$ to form uniform stable microemulsions with droplet sizes between 36-70 nm. $PhB_{pH6.5}$-$SEDDS_{Migylol812}$ and $PhB_{pH6.5}$-$SEDDS_{OliveOil}$ displayed significantly different mean droplet sizes (* $p < 0.05$) (Table 2). Droplet sizes of FaSSIF-$SEDDS_{Migylol812}$ and FaSSIF-$SEDDS_{OliveOil}$ also differed (* $p < 0.05$) and were smaller than $PhB_{pH6.5}$-$SEDDS_{Migylol812}$ and $PhB_{pH6.5}$-$SEDDS_{OliveOil}$. All PDI's obtained were below 0.26 indicating droplet sizes on dispersion were moderately homogenous. In terms of charge, values close to zero mV were observed for $PhB_{pH6.5}$-$SEDDS_{Migylol812}$ and $PhB_{pH6.5}$-$SEDDS_{OliveOil}$ as all SEDDS excipients were non-ionic and neutral. FaSSIF displayed an overall net negative charge (-14.67 mV), which remained, though reduced in magnitude, through dispersion of $SEDDS_{OliveOil}$ (-5.73 mV) and $SEDDS_{Migylol812}$ (-5.35 mV) (Table 2).

Solubility in Biorelevant SEDDS Dispersions – Comparison of $SEDDS_{Migylol812}$ and $SEDDS_{OliveOil}$

For the 30 drugs, solubility in FaSSIF-$SEDDS_{Migylol812}$ was higher than FaSSIF-$SEDDS_{OliveOil}$ as a paired bootstrap test revealed a significant difference in drug solubility between these medium chain and long chain lipid dispersions (* $p < 0.05$). Comparatively, the beta coefficient of the regression line for FaSSIF-$SEDDS_{Migylol812}$ versus FaSSIF-$SEDDS_{OliveOil}$ was significant according to a bootstrap for coefficients test (* $p < 0.05$). A strong correlation was established between drug solubility in FaSSIF-$SEDDS_{Migylol812}$ and FaSSIF-$SEDDS_{OliveOil}$ (r^2 0.97) (Figure 1), suggesting that for every 100 unit increase in FaSSIF-$SEDDS_{OliveOil}$ solubility units, FaSSIF-$SEDDS_{Migylol812}$ increases on average by 105.6 solubility units. Consequentially, this indicates that solubility determined in one lipid dispersion may be used to estimate solubility in the other.

Solubility Ratio Trends

Solubility ratios (SR) for 30 PWSDs upon dispersion of two SEDDS was experimentally determined (Figure 2), where SR >1 was seen in all cases, indicative of increased drug solubility on SEDDS dispersion in intestinal media. SRs ranged from 1.13 - 64.4 fold for $SEDDS_{Migylol812}$ and from 1.04-59.7 fold for $SEDDS_{OliveOil}$. In presence of both SEDDS, Clotrimazole and Fenofibrate displayed the highest SRs. Trends in ionisable drugs were analysed. Cationic drugs appeared to consistently display high SR, with all such compounds displaying solubility gains of >2, with 3 and 2 drugs respectively displaying SR >10 fold in presence of $SEDDS_{Migylol812}$ and $SEDDS_{OliveOil}$. In contrast, solubility gains for anionic compounds appeared less pronounced, with 8/9 anionic drugs displayed SR <5. However, for Candesartan Cilexetil, a SR >16 was observed with both SEDDS. Candesartan Cilexetil is an ampholyte where the hydrogen attached to the O-$CH(CH_3)$-O group in the cilexetil side chain is moderately acidic, being between the oxygen rich ester moieties, while also possessing a basic functional group. This ampholytic nature may have contributed to its deviation from the general trends observed for other anionic drugs. Neutral drugs displayed a wide range of SRs, while Celecoxib and Venetoclax deviated strongly from the trend of similar SRs in $SEDDS_{Migylol812}$ and $SEDDS_{OliveOil}$, with Celecoxib displaying a SR_{MC} of 17 compared to a SR_{LC} of 7, while Venetoclax also displayed a

difference between SR_{MC} and SR_{LC} i.e. 12 versus 7. To assess SEDDS ability to mirror solubility increases in fed-state versus and fasted-state media, SRs obtained were compared to FeSSIF/FaSSIF solubility ratios. SR_{MC} and SR_{LC} exceeded $SR_{FeSSIF/FaSSIF}$ for 24 and 23 of the 30 drugs respectively (Figure 3). This observation confirms the utility of SEDDS as effective bio-enabling systems to bridge the fasted-fed state solubility gap (35).

Computational Prediction of Biorelevant Solubility Gain with SEDDS.

Linear Regression revealed weak correlations between both SR_{LC} and SR_{MC} versus individual drug properties. Lipophilicity and T_m, commonly utilised as guides towards LBF suitability, displayed poor quantitative relationships e.g. logP (r^2 0.33, 0.32), $logD_{6.5}$ (r^2 0.43, 0.35) and T_m (r^2 0.23, 0.25). Therefore, a combination of variables were required to improve quantitative prediction accuracy. Firstly, PCA verified the dataset structural diversity (Supplementary Materials). PLS model development resulted in predictive PLS models for both SRs ($logSR_{MC}$ and $logSR_{LC}$). The PLS models used 1-2 principle components (PC) and 5-6 variables. The $logSR_{MC}$ 1 PC model produced predictions of r^2 calibration 0.81, r^2 validation 0.73 requiring 5 variables; $logD_{6.5}$, melting point (T_m), molecular weight (MW), aromatic bonds as fraction of total bonds (F_AromB) and Atom-Type Cumulative Electrotopological State (E-state) index for methylene carbons (SssCH2). While the $logSR_{LC}$ 2 PC model required 6 variables; $LogD_{6.5}$, MW, T_m, F_AromB, SssCH2 and number of aliphatic rings (N_AliphR) to produce predictions of r^2 calibration 0.77, r^2 validation 0.67. These models demonstrated good predictions of test sets, summarized in Table 3.

Enhanced Biorelevant Solubility Ratio Equation

As 5-6 descriptors could predict SR_{MC} and SR_{LC}, multiple linear regression was performed to produce easily interpretable predictive equations. All significant variables from PLS modelling were initially included in MLR. Insignificant variables ($p > 0.05$) from these initial equations were subsequently removed, resulting in final equations with higher F-values and significant variables. Two equations were produced (Table 3), both utilising 3 properties: $logD_{6.5}$, T_m and F_AromB, Similarities between equations was expected due to the high correlation between dispersed SEDDS (Figure 1).

Use of Predicted Solubility Ratios to Predict Drug DCS Class with SEDDS.

Application of the equations to predict drug DCS class with SEDDS was assessed and accuracy compared to comparative DCS classifications using experimentally determined solubility's upon SEDDS dispersion. DCS Permeability classifications were estimated using drug permeability predictions from the ADMET Predictor 9.5 (Simulations Plus, USA). While use of a computationally derived permeability estimate has been applied in other studies (36), it must be acknowledged that drug specific effects may not be adequately captured in these predicted permeability estimates. In total, using experimental solubility's, 10 drugs overcame a solubility limitation i.e. transitioned to DCS Class I/III. Using the $Do_{(Predicted)}$ approach (Equation 6), this transition was correctly predicted for 8/10 drugs (Table 4) i.e. Clotrimazole, Cinnarizine, Fenofibrate, Isotretinoin, Naproxen, Terfenadine, Glipizide and Venetoclax. DCS Classification using $Do_{(SEDDS)}$ (Equation 4) also resulted in transitions to "good solubility" for Candesartan Cilexetil and Celecoxib ($SEDDS_{Migylol812}$ only), however, as previously discussed experimental results for both drugs differed significantly from general trends

observed, which may suggest a drug specific effects in these cases that was not captured in the MLR equations.

Predicted DCS Classifications of Commercial LBF Drugs.

The utility of the $Do_{(Predicted)}$ approach to guide a LBF formulation strategy was subsequently assessed by applying the MLR equations to a range of drugs that have been successfully licensed as LBFs. A total of 49 drugs were selected initially, and the DCS classification using FaSSIF solubility alone was employed to determine DCS class. In total, 23 drugs were initially classified as DCS class I/III, and therefore, did not display solubility limitations. These compounds were therefore excluded from further analysis as a bio-enabling strategy was not considered necessary. Applying the $Do_{(Predicted)}$ approach to the remaining 26 drugs, 10 drugs were predicted to transition from poor to good solubility, and a further 7 drugs were found to transition from DCS Class IIb to Class IIa i.e "dissolution rate limited" which can offer delivery opportunities, where the compensatory influence of high permeability has been stated to be significant for acceptable oral absorption during the transit time in the intestine (22). Therefore, this approach predicted that in 65.4% (i.e. 17/26) of drugs, a SEDDS approach was likely to overcome solubility limited absorption. Of the 9 drugs that remained in poor solubility classification after applying the $Do_{(Predicted)}$ approach, 8 were DCS Class IV, which may indicate that decisions to employ a SEDDS approach were not solely influenced by solubility considerations and that other factors, such as increased permeability, may have been a consideration in the choice to develop as a LBF (Table 5).

Discussion

Over the last two decades, significant strides have been made in applying computational approaches across the full spectrum of drug development (37). In their many forms, computational tools can include discovering new lead candidates with optimal drug-receptor binding affinity (e.g. Quantitative Structural Activity Relationships (QSAR)), to guiding on optimal physiochemical profiles (e.g. Quantitative Structural Property Relationships (QSPR)) or predictions of biopharmaceutical properties including solubility and permeability (9). While the major advances in the use of computational tools to-date have been focused on chemical structural design to assist the selection of new drug substances with optimal pharmacodynamic and/or pharmacokinetic properties, commonly referred to as "druggability", more recently, the use of computational tools to guide on formulation design, or computational pharmaceutics, have been reported (9, 38, 39). These include approaches such as computational biopharmaceutical drug profiling, recently reported as an approach to predict physiochemical and molecular properties of drug candidates that render them more or less suitable for formulation via a specific bio-enabling formulation approach (9). Accordingly, there exists an increasing focus on development of reliable computational pharmaceutics tools, capable of guiding selection of appropriate bio-enabling formulation strategies, in particular for drug candidates which display either solubility and/or permeability limitations.

LBFs are one such bio-enabling formulation technology that exploit the benefit of lipid excipients to harness the absorption pathways of dietary fats, leading to increased intestinal drug solubility and

improving intestinal absorption. The benefits of lipid excipients to increase drug solubility were clearly prevalent in this study, where increased solubility was observed for all 30 PWSDs following dispersion of the SEDDS in biorelevant media. Indeed, the solubility increases observed were on average higher than the fed:fasted biorelevant solubility ratio, as SR_{MC} and SR_{LC} exceeded $SR_{FeSSIF/FaSSIF}$ for 24 and 23 of the 30 drugs respectively (Figure 3). However, despite clear benefits as a bio-enabling technology, it is generally considered that LBFs have an unfulfilled potential in a commercial sense. Over the last decade, prevalence of commercial LBFs appears to be decreasing relative to Solid Dispersions (SD) (4), reflecting improved scientific knowledge on the pharmaceutical benefits of SDs in terms of bio-enabling effects (e.g. increased drug solubility), but also an improved understanding of factors influencing industrial scalability and regulatory approval (e.g. long term stability). On the other hand, the prevalence of commercial LBFs has tended to be relatively few, reflecting gaps in understanding both in terms of bio-enabling benefits and from an industrial perspective, as recently reviewed (5). With this in mind, significant strides have been made in the use of *in silico* approaches to reliably predict dose loading capacity in LBFs (11, 12). This current study sought to advance the application of computational pharmaceutics tools to consider the impact of *in vivo* dispersion of SEDDS on drug solubility in GI fluids. In recognition of the importance of *in vivo* dispersion on SEDDS performance (40), we hypothesised that computational prediction of drug solubility increases seen upon dispersion of SEDDS in simulated biorelevant fluids is likely to be a key performance indicator of whether a SEDDS approach can effectively solubilise the dose *in vivo*. As such, a computationally predicted solubility ratio (SR), based on drug properties in combination with experimentally determined solubility in FaSSIF, would support more informed decisions on formulation options in early development, by allowing estimation of a biopharmaceutically relevant Do.

Resultantly, our hypothesis that a relationship could be elucidated between a biorelevant SR for a SEDDS formulation and drug properties was demonstrated and shown to be robust. We observed, on a dataset of 30 PWSDs using PLS computational modelling, that 5-6 drug properties were sufficient to reliably predict SR upon dispersion of two prototype SEDDS ($logSR_{MC}$ r^2 0.81, $logSR_{LC}$ r^2 0.77). Subsequently, employment of MLR facilitated simplified equations for SR to be generated, requiring only 3 drug properties namely, partition coefficient pH 6.5 ($logD_{6.5}$), melting point (T_m) and aromatic bonds as fraction of total bonds (FArom_B). These represent common drug properties typically identified and integrated into an early stage pharmaceutical drug profiling environment (10), forgoing requirements for molecular fragment profiling or specialised chemometric software.

Inclusion of drug properties in this computational model, implies their importance to SR upon SEDDS dispersion at a mechanistic level. For the $logSR_{MC}$ model, important descriptors were $LogD_{6.5}$, T_m, MW, F_AromB and SssCH2. Additionally, the $logSR_{LC}$ model also included N_AlipR. In terms of, $logD_{6.5}$, T_m and MW, these are widely recognised drug properties from a pharmaceutical profiling context. In both PLS models, $logD_{6.5}$ and MW were positively correlated with SR while T_m was negatively correlated. Inclusion of a partition coefficient descriptor was not unexpected due to addition of lipophilic SEDDS to the media, while $logD_{6.5}$ was previously observed to be strongly correlated with PWSD solubilisation in biorelevant media (25), and an influential descriptor in modelling the

FaSSIF/PhB$_{pH6.5}$ ratio (24). Additionally, distribution coefficient has been used to characterise drug release from SEDDS, or more specifically the drug diffusion process from the SEDDS pre-concentrate into aqueous media has been related to logD$_{SEDDS/RM}$ i.e. the distribution coefficient of solubility in SEDDS pre-concentrate and the release medium (41). Conversely, the negative correlation between T_m and SR is most likely attributable to high T_m molecules exhibiting solid state limited solubility or 'brick dust' drugs, which results in poor solubility in lipid excipients, translating to more modest SR values upon SEDDS dispersion. While the importance of MW as a descriptor is not unexpected given the influence of MW on both crystalline structure characteristics and solvation properties, in contrast to trends observed between MW and aqueous solubility (42), MW and SR in this case are positively correlated. Accordingly, as increasing size negatively influences aqueous solubility, MW may be indirectly conveying information regarding relative drug affinity for lipophilic formulation excipients to that for the comparatively more aqueous environment within the biorelevant medium. Finally, a recent retrospective analysis of selected physiochemical and molecular properties of drugs produced commercially as LBF products versus commercial SD drugs and a database of drugs not produced via either bio-enabling approach, found logD, T_m and MW to be significant descriptors signally commercial success with LBFs. Similar to this study, increasing logD and MW were found to be significant for LBF commercial success, while a lower relative drug T_m was found to be significant to reach commercialisation (4). The fact that these descriptors were significant in both a retrospective analysis of successfully commercialised LBFs and in this prospective SR prediction upon SEDDS dispersion, re-emphasises their importance as contributing factors to drug-LBF technology success and suitability.

Additionally, SssCH2, F_AromB and N_AlipR were significant in PLS modelling. F_AromB was positively correlated to SR. While this positive correlation is in contrast to previous predictions of HIF solubility (28), it is likely that as increasing aromatic ring count decreases aqueous solubility (43, 44), and an increase in affinity for lipid excipients is seen. In this case, compounds with larger aromatic structures are likely to have a negative influence on aqueous solubility. Upon SEDDS dispersion, such compounds will associate with greater affinity to the lipid rich microemulsions droplets formed, resulting in a higher SR. However, contributions of aromaticity are likely complex, reliant on numerous factors including; attached substituents and their polarity, existence of 'through resonance' with attached substituents, as well as ion-dipole and dipole-dipole interactions with other moieties. Number of aromatic bonds was previously significant for *in silico* prediction of FeSSIF/FeSSIF blank buffer, further highlighting the significance of aromaticity for solubility in media with increasing lipids (25). N_AlipR also influences drug shape and size and is also affected by adjacent moieties. Meanwhile, SssCH2 examines the topological and electronic features of a structure (45) and was previously significant in an *in silico* prediction of solubility in FaSSIF buffer (25).

This work also investigated other factors influencing SR in order to understand of how drugs associate with biorelevant SEDDS dispersions. In terms of drug ionisation, general trends of higher SR for cationic (charged basic) versus anionic (charged acidic) drugs were observed. This observations are in line with previous research where solubility increases in biorelevant media versus

corresponding blank buffers for bases and neutral drugs were higher than acids (25). Such increases for cationic drugs, have previously been suggested to stem from favourable electrostatic interactions between negatively charged polar head groups of taurocholate bile salts (46, 47) and positively charged drugs. In this case, such bile salt related electrostatic interactions are likely to occur in both FaSSIF and FaSSIF-SEDDS, with net negative charges observed for all three media. The general trend for increased SR for cationic compounds occurred despite an overall reduction in net negative charge in both FaSSIF-SEDDS media relative to FaSSIF, demonstrating the possibility that additional electrostatic interactions may exist. As both alterations to droplet sizes and to overall charge of the media upon SEDDS dispersion in FaSSIF versus $PhB_{pH6.5}$ were observed, interactions between the SEDDS and biorelevant solubilising components of FaSSIF are probable. In particular, the negative charges of FaSSIF-$SEDDS_{Migylol812}$ (-5.35 mV) and FaSSIF-$SEDDS_{OliveOil}$ (-5.73 mV), were intermediate of the overall charges of FaSSIF (-14.67 mV) and the values close to zero observed upon SEDDS dispersion in $PhB_{pH6.5}$ (-0.76 mV, -1.27 mV), suggesting surface association of charged bile salts to the oil droplets formed upon SEDDS dispersion. Such an association was previously proposed upon initial *in vitro* dispersion of a SEDDS in a biorelevant media (48). It therefore could be suggested upon SEDDS dispersion, favourable interactions between cationic drugs and these charged bile salts found at the oil droplet surface may help explain the increased SRs observed. Previously, electrostatic interactions between cationic drugs and free fatty acids in post digestive media have also been suggested as a potential mechanism for increased drug solubilisation.(40). However, presently such interactions are poorly understood and electrostatic interactions appear to not be the sole solubilising mechanism involved, given that both neutral and cationic drugs also displayed SRs between 1.1 and 51, hence indicating that there are a number of additional factors governing drug associated with mixed colloidal dispersion.

In terms of excipient effects, $SEDDS_{Migylol812}$ and $SEDDS_{OliveOil}$ were compared. A strong correlation was observed between solubility in FaSSIF-$SEDDS_{Migylol812}$ versus FaSSIF-$SEDDS_{OliveOil}$, suggesting that strong correlations previously observed between drug solubility in MCT versus LCT preconcentrates, and C8 versus C10 triglycerides are also observed upon SEDDS dispersion of these exemplary MCTs and LCTs (11, 15). In all cases solubility in FaSSIF-$SEDDS_{Migylol812}$ was higher than FaSSIF-$SEDDS_{OliveOil}$, with an overall significant difference observed (* $p < 0.05$). However, the extent of solubility difference between both was relatively small i.e. for 20 out of the 30 drugs the difference was < 20 %. Therefore, in terms of the choice of these exemplary MCT or LCT containing SEDDS, the practical implications in terms of solubility difference on dispersion in biorelevant buffer appear relatively minor. The merits of MCT versus LCT have been widely discussed (49, 50). While in general, drug solubility in most examples of MCTs is higher (1, 51), following formulation digestion, the digestion products of LCT may confer additional advantages (52), while it must also be acknowledged that these trends may not be observed for all MCT and LCTs comparisons. This study also identified two specific drug examples, namely Celecoxib and Venetoclax, where large differences in SR were observed, relating to large solubility percentage differences (58% and 43%) being observed between both SEDDS dispersions. The possible reason for these higher associations with dispersed $SEDDS_{Migylol812}$ for these two neutral drugs are unclear, however this highlights a potential

limitation of computational predictions to capture specific drug-excipient solubility effects. Therefore, future work with a wider range of drugs could help to increase robustness of the predictions achieved.

Overall, this work endeavoured to advance the field of computational pharmaceutics by demonstrating the capacity for such predictive tools to inform developability, and specifically to guide formulation decisions regarding SEDDS by assessing their ability to improve the biopharmaceutical dose number. $Do_{(Predicted)}$ (Equation 6) can be easily applied as a computational pharmaceutics tool to guide formulation suitability, requiring only 3 readily obtainable drug properties, in addition to an experimentally determined drug solubility in FaSSIF. The suitability of $Do_{(Predicted)}$ to forecast developability was validated by comparing predicted to experimental Do values, showing that 8/10 drugs were correctly predicted to transition to a “good solubility” DCS class (I/III). The two drugs, Candesartan Cilexetil and Celecoxib that were not predicted to transition most likely reflect the limitation of the model to capture drug specific solubility increases, as discussed previously. Subsequently, to demonstrate the real-time applicability of such predictions in a pharmaceutical developability context, $Do_{(Predicted)}$ was applied to a drug dataset outside the training and test sets, namely drugs previously successfully produced as commercial LBF products. The $Do_{(Predicted)}$ approach predicted that two out of three (65.4%) of these drugs would offer benefits for development as a LBF. Furthermore, when DCS classes using FaSSIF solubility versus DCS class using predicted solubility with SEDDS were compared, 8 of the 9 commercial drugs which demonstrated no class transition were DCS Class IV. Therefore, as these predictions are based upon drug solubility gains with SEDDS it is likely that permeability considerations, not only solubility benefits, were influential in the development of these poorly soluble and poorly permeable drugs with LBFs.

Comparable to the stated limitation of the original DCS classification system (22), potential for supersaturation was not explored in these predictions. This would have particular relevance for ionisable drugs displaying pH dependent solubility, while weakly basic drugs in particular exhibit higher solubility in gastric media, along with potential for intestinal supersaturation and precipitation. Further limitations of the predictions are also acknowledged in terms of the deliberate omission of exploration of the effect of SEDDS digestion on drug solubility. We therefore acknowledge that this tool is conservative in its approach to solubility predictions and the solubility gains are likely to be under predictive of the kinetic solubility’s achieved in the gastrointestinal tract. However, from an industry perspective, where conservative risk:benefit approaches are often applied to formulation development, this low risk approach may be in line with current industrial preferences. To overcome any conservative nature in the application of a predicted Do, we suggest incorporation of this tool into the refined Developability Classification System (rDCS) as part of the initial “standardised investigations” (23). For a weakly basic drug, customised investigations such as the small-scale supersaturation/precipitation experiments as specified in the rDCS could be then triggered to test the potential effects of supersaturation.

Conclusion

Through combinations of *in silico* predictions based on drug properties, and drug solubility screening in FaSSIF, this work demonstrated capacity for computational pharmaceutics to inform drug developability. By applying a computational pharmaceutics approach this study identified drug properties that can be used to predict SR for SEDDS dispersions. The results demonstrated that integration of biorelevant experimentally determined FaSSIF solubility into computationally predicted dose numbers (i.e. combining molecular, physiochemical and biopharmaceutical properties), allows more reliable biopharmaceutically relevant and data-driven decisions to be made on drug-SEDDS developability. While it is acknowledged that *in silico* predictions are not intended to completely circumvent experimental solubility screening, when used in conjunction with appropriate screening assays, such tools can guide likely successful bio-enabling approaches in a biopharmaceutically informed manner. In order to advance this growing field of computational pharmaceutics for LBFs, renewed emphasis should be placed upon creating validated and increasingly robust computational predictions of drug developability with bio-enabling formulations.

Acknowledgments

This work was supported under funding from the Irish Research Council Post Graduate Scholarship Project number GOIPG/2018/883. The authors would like to thank Ana Luiza Queiroz for her help with PLS modelling as well as Dr. John Quigley for his greatly valued guidance on chemistry aspects of the analysis.

1. Feeney OM, Crum MF, McEvoy CL, Trevaskis NL, Williams HD, Pouton CW, et al. 50years of oral lipid-based formulations: Provenance, progress and future perspectives. Adv Drug Deliv Rev. 2016;101:167-94.
2. Ditzinger F, Price DJ, Ilie AR, Kohl NJ, Jankovic S, Tsakiridou G, et al. Lipophilicity and hydrophobicity considerations in bio-enabling oral formulations approaches - a PEARRL review. J Pharm Pharmacol. 2019;71(4):464-82.
3. Strickley RG. Currently Marketed Oral Lipid-Based Dosage Forms. In: Hauss DJ, editor. Oral Lipid-Based Formulations Enhancing the Bioavailability of Poorly Water-Soluble Drugs. 170. Drugs and Pharmaceutical Sciences: Taylor & Francis Group; 2007. p. 1-32.
4. Bennett-Lenane H, O'Shea JP, O'Driscoll CM, Griffin BT. A Retrospective Biopharmaceutical Analysis of >800 Approved Oral Drug Products: Are Drug Properties of Solid Dispersions and Lipid-Based Formulations Distinctive? Journal of Pharmaceutical Sciences. 2020;109(11):3248-61.
5. Holm R. Bridging the gaps between academic research and industrial product developments of lipid-based formulations. Advanced Drug Delivery Reviews. 2019;142:118-27.
6. O'Driscoll CM, Griffin BT. Biopharmaceutical challenges associated with drugs with low aqueous solubility--the potential impact of lipid-based formulations. Adv Drug Deliv Rev. 2008;60(6):617-24.
7. Kuentz M. Lipid-based formulations for oral delivery of lipophilic drugs. Drug Discovery Today: Technologies. 2012;9(2):e97-e104.
8. Chen XQ, Gudmundsson OS, Hageman MJ. Application of lipid-based formulations in drug discovery. J Med Chem. 2012;55(18):7945-56.

9. Bergstrom CAS, Charman WN, Porter CJH. Computational prediction of formulation strategies for beyond-rule-of-5 compounds. Adv Drug Deliv Rev. 2016;101:6-21.
10. Bergström CA, Holm R, Jørgensen SA, Andersson SB, Artursson P, Beato S, et al. Early pharmaceutical profiling to predict oral drug absorption: current status and unmet needs. Eur J Pharm Sci. 2014;57:173-99.
11. Persson LC, Porter CJ, Charman WN, Bergstrom CA. Computational prediction of drug solubility in lipid based formulation excipients. Pharm Res. 2013;30(12):3225-37.
12. Alskar LC, Porter CJ, Bergstrom CA. Tools for Early Prediction of Drug Loading in Lipid-Based Formulations. Mol Pharm. 2016;13(1):251-61.
13. Sacchetti M, Nejati E. Prediction of drug solubility in lipid mixtures from the individual ingredients. AAPS PharmSciTech. 2012;13(4):1103-9.
14. Cao Y, Marra M, Anderson BD. Predictive relationships for the effects of triglyceride ester concentration and water uptake on solubility and partitioning of small molecules into lipid vehicles. J Pharm Sci. 2004;93(11):2768-79.
15. Gautschi N, Bergstrom CA, Kuentz M. Rapid determination of drug solubilization versus supersaturation in natural and digested lipids. Int J Pharm. 2016;513(1-2):164-74.
16. Boyd BJ, Bergström CAS, Vinarov Z, Kuentz M, Brouwers J, Augustijns P, et al. Successful oral delivery of poorly water-soluble drugs both depends on the intraluminal behavior of drugs and of appropriate advanced drug delivery systems. European Journal of Pharmaceutical Sciences. 2019;137:104967.
17. Khan J, Rades T, Boyd B. The Precipitation Behavior of Poorly Water-Soluble Drugs with an Emphasis on the Digestion of Lipid Based Formulations. Pharm Res. 2016;33(3):548-62.
18. Bevernage J, Brouwers J, Annaert P, Augustijns P. Drug precipitation-permeation interplay: supersaturation in an absorptive environment. Eur J Pharm Biopharm. 2012;82(2):424-8.
19. Williams HD, Sassene P, Kleberg K, Calderone M, Igonin A, Jule E, et al. Toward the establishment of standardized in vitro tests for lipid-based formulations, part 4: proposing a new lipid formulation performance classification system. J Pharm Sci. 2014;103(8):2441-55.
20. Klein S. The use of biorelevant dissolution media to forecast the in vivo performance of a drug. Aaps j. 2010;12(3):397-406.
21. Schwebel HJ, van Hoogevest P, Leigh ML, Kuentz M. The apparent solubilizing capacity of simulated intestinal fluids for poorly water-soluble drugs. Pharm Dev Technol. 2011;16(3):278-86.
22. Butler JM, Dressman JB. The developability classification system: application of biopharmaceutics concepts to formulation development. J Pharm Sci. 2010;99(12):4940-54.
23. Rosenberger J, Butler J, Dressman J. A Refined Developability Classification System. J Pharm Sci. 2018;107(8):2020-32.
24. Fagerberg JH, Al-Tikriti Y, Ragnarsson G, Bergstrom CA. Ethanol effects on apparent solubility of poorly soluble drugs in simulated intestinal fluid. Mol Pharm. 2012;9(7):1942-52.
25. Fagerberg JH, Tsinman O, Sun N, Tsinman K, Avdeef A, Bergstrom CA. Dissolution rate and apparent solubility of poorly soluble drugs in biorelevant dissolution media. Mol Pharm. 2010;7(5):1419-30.
26. Van den Bergh A, Van Hemelryck S, Bevernage J, Van Peer A, Brewster M, Mackie C, et al. Preclinical Bioavailability Strategy for Decisions on Clinical Drug Formulation Development: An In Depth Analysis. Mol Pharm. 2018;15(7):2633-45.
27. Griffin BT, Kuentz M, Vertzoni M, Kostewicz ES, Fei Y, Faisal W, et al. Comparison of in vitro tests at various levels of complexity for the prediction of in vivo performance of lipid-based formulations: case studies with fenofibrate. Eur J Pharm Biopharm. 2014;86(3):427-37.
28. Fagerberg JH, Karlsson E Fau - Ulander J, Ulander J Fau - Hanisch G, Hanisch G Fau - Bergstrom CAS, Bergstrom CA. Computational prediction of drug solubility in fasted simulated and aspirated human intestinal fluid. (1573-904X (Electronic)).
29. Klumpp L, Leigh M, Dressman J. Dissolution behavior of various drugs in different FaSSIF versions. European Journal of Pharmaceutical Sciences. 2020;142:105138.

30. Augustijns P, Wuyts B, Hens B, Annaert P, Butler J, Brouwers J. A review of drug solubility in human intestinal fluids: implications for the prediction of oral absorption. Eur J Pharm Sci. 2014;57:322-32.
31. Wiest J, Saedtler M, Bottcher B, Grune M, Reggane M, Galli B, et al. Geometrical and Structural Dynamics of Imatinib within Biorelevant Colloids. Mol Pharm. 2018;15(10):4470-80.
32. Bhattacharjee S. DLS and zeta potential - What they are and what they are not? J Control Release. 2016;235:337-51.
33. Fagerberg JH, Bergstrom CA. Intestinal solubility and absorption of poorly water soluble compounds: predictions, challenges and solutions. Ther Deliv. 2015;6(8):935-59.
34. Forina M, Lanteri S, Oliveros MCC, Millan CP. Selection of useful predictors in multivariate calibration. Analytical and Bioanalytical Chemistry. 2004;380(3):397-418.
35. O'Shea JP, Holm R, O'Driscoll CM, Griffin BT. Food for thought: formulating away the food effect - a PEARRL review. J Pharm Pharmacol. 2019;71(4):510-35.
36. Hosey CM, Benet LZ. Predicting the extent of metabolism using in vitro permeability rate measurements and in silico permeability rate predictions. Mol Pharm. 2015;12(5):1456-66.
37. Kumar N, Hendriks BS, Janes KA, de Graaf D, Lauffenburger DA. Applying computational modeling to drug discovery and development. Drug Discovery Today. 2006;11(17):806-11.
38. Hossain S, Kabedev A, Parrow A, Bergström CAS, Larsson P. Molecular simulation as a computational pharmaceutics tool to predict drug solubility, solubilization processes and partitioning. Eur J Pharm Biopharm. 2019;137:46-55.
39. Mehta CH, Narayan R, Nayak UY. Computational modeling for formulation design. Drug Discovery Today. 2019;24(3):781-8.
40. Alskar LC, Keemink J, Johannesson J, Porter CJH, Bergstrom CAS. Impact of Drug Physicochemical Properties on Lipolysis-Triggered Drug Supersaturation and Precipitation from Lipid-Based Formulations. Mol Pharm. 2018;15(10):4733-44.
41. Bernkop-Schnürch A, Jalil A. Do drug release studies from SEDDS make any sense? J Control Release. 2018;271:55-9.
42. Tolls J, van Dijk J, Verbruggen EJM, Hermens JLM, Loeprecht B, Schüürmann G. Aqueous Solubility−Molecular Size Relationships: A Mechanistic Case Study Using C10- to C19-Alkanes. The Journal of Physical Chemistry A. 2002;106(11):2760-5.
43. Ritchie TJ, Macdonald SJ. The impact of aromatic ring count on compound developability--are too many aromatic rings a liability in drug design? Drug Discov Today. 2009;14(21-22):1011-20.
44. Wassvik CM, Holmén AG, Draheim R, Artursson P, Bergström CAS. Molecular Characteristics for Solid-State Limited Solubility. Journal of Medicinal Chemistry. 2008;51(10):3035-9.
45. Kier LB, Hall LH. An Electrotopological-State Index for Atoms in Molecules. Pharmaceutical Research. 1990;7(8):801-7.
46. Clulow AJ, Parrow A, Hawley A, Khan J, Pham AC, Larsson P, et al. Characterization of Solubilizing Nanoaggregates Present in Different Versions of Simulated Intestinal Fluid. J Phys Chem B. 2017;121(48):10869-81.
47. Wickham M, Garrood M, Leney J, Wilson PD, Fillery-Travis A. Modification of a phospholipid stabilized emulsion interface by bile salt: effect on pancreatic lipase activity. J Lipid Res. 1998;39(3):623-32.
48. Fatouros DG, Bergenstahl B, Mullertz A. Morphological observations on a lipid-based drug delivery system during in vitro digestion. Eur J Pharm Sci. 2007;31(2):85-94.
49. Pouton CW, Porter CJH. Formulation of lipid-based delivery systems for oral administration: Materials, methods and strategies. Advanced Drug Delivery Reviews. 2008;60(6):625-37.
50. Christensen J, Schultz K, Mollgaard B, Kristensen HG, Mullertz A. Solubilisation of poorly water-soluble drugs during in vitro lipolysis of medium- and long-chain triacylglycerols. Eur J Pharm Sci. 2004;23(3):287-96.
51. Mishra M. Vegetable Oil-based Formulations for Controlled Drug Delivery. Handbook of Encapsulation and Controlled Release: CRC Press; 2015. p. 1383.

52. Williams HD, Sassene P, Kleberg K, Calderone M, Igonin A, Jule E, et al. Toward the establishment of standardized in vitro tests for lipid-based formulations, part 3: understanding supersaturation versus precipitation potential during the in vitro digestion of type I, II, IIIA, IIIB and IV lipid-based formulations. Pharm Res. 2013;30(12):3059-76.

Table 1:

Drug Compound	MW (g/mol)	clogP	$logD_{6.5}$	Acid/Base /Neutral	pKa (% ionised at pH 6.5)	Classification of Charge pH 6.5	T_m (°C)	PSA ($Å^2$)	HBD	HBA	Ro5	Max Dose (mg)
Albendazole	265.3	2.81	2.80	Ampholyte	10.26 (0%), 2.8 (0%)	0	209	67.01	2	3	0	200
Candesartan Cilexetil	610.7	5.70	2.89	Ampholyte	6 (76%)	-	163	143.3	1	8	1	32
Carbamazepine	236.27	2.40	2.40	Basic	13.9 (0%)	0	190.2	46.3	1	1	0	300
Carvedilol	406.4	3.88	2.36	Basic	7.8 (95%)	+	114.5	75.7	3	5	0	25
Celecoxib	381.37	3.81	3.81	Acidic	11.1 (0%)	0	158	77.98	1	3	0	200
Cinnarizine	368.6	4.92	3.98	Basic	8.4 (99%)	+	119	6.48	0	2	1	25
Clofazimine	473.40	7.11	4.54	Basic	8.51 (99%)	+	211	40	1	4	1	50
Clotrimazole	344.9	5.08	5.06	Basic	6.7 (96%)	+	142	17.8	0	1	1	10
Danazol	337.5	4.26	4.26	Neutral	-	0	227	46.3	1	2	0	200
Dipyridamole	504.64	3.10	3.02	Basic	6.59 (55%)	+	163	145	4	12	2	200
Felodipine	384.3	5.03	5.03	Basic	5.07 (3%)	0	143	64.4	1	3	0	10
Fenofibrate	360.9	5.20	5.20	Neutral	-	0	79	52.6	0	4	1	150
Glipizide	445.5	2.12	1.48	Acidic	5.9 (80%)	-	201.5	130.15	3	6	0	10
Griseofulvin	352.77	2.51	2.51	Neutral	-	0	220	71.1	0	6	0	500
Haloperidol	375.9	3.82	2.06	Basic	8.3 (98%)	+	151	40.5	1	3	0	20
Indomethacin	357.8	4.03	1.45	Acidic	4.5 (99%)	-	160	68.5	1	4	0	50
Irbesartan	428.53	3.68	2.84	Ampholyte	4.12 (0%), 7.4 (11%)	0	180.5	87.13	1	5	0	300
Isotretinoin	300.44	6.07	3.99	Acidic	4 (99%)	-	174	37.3	1	2	1	40
Itraconazole	705.7	4.89	4.89	Basic	3.7 (0%)	0	166	104.7	0	9	2	100
Ketoconazole	531.43	3.67	3.51	Basic	6.75 (64%), 4.22 (0%)	+	146	69.06	0	6	1	200
Mefenamic Acid	241.29	4.91	2.36	Acidic	3.89 (99%)	-	230.5	49.3	2	3	1	500
Naproxen	230.26	3.21	1.10	Acidic	4.15 (99%)	-	153	46.5	1	3	0	500
Nifedipine	346.34	3.10	3.10	Acidic	3.93 (99%)	-	173	110.45	1	5	0	90
Phenytoin	252.27	2.09	2.07	Acidic	8.3 (2%)	0	296.5	58.2	2	2	0	300
Progesterone	314.5	3.94	3.94	Neutral	-	0	128	34.1	0	2	0	200
Spironolactone	416.57	3.28	3.28	Neutral	-	0	134.5	60.44	0	3	0	100
Tamoxifen	371.52	6.59	4.61	Basic	8.5 (99%)	+	97	12.5	0	2	1	20
Terfenadine	471.67	5.60	3.61	Basic	10 (99%)	+	147	47.3	2	3	1	60
Tolfenamic Acid	261.7	5.13	2.44	Acidic	5.11(96%)	-	213	49.3	2	3	0	200
Venetoclax	868.44	6.76	6.54	Ampholyte	3.4 (99%), 10.3 (99%)	0	138	172.03	3	10	2	100

Table 1: Selection of Physiochemical and Molecular Properties of Investigated Compounds collated from literature or ADMET Predictor 9.5. 0 = no charge at pH 6.5, + = positive charge, - = negative charge. Am = Ampholyte. % refers to the percentage of the drug's ionisable groups ionised at pH 6.5 according to the Henderson-Hasselbalch Equation.

Table 2:

Media	Size nm (SD)	PDI (SD)	ζ-potential mV (SD)
$PhB_{pH6.5}$-SEDDS *Migylol812*	47.71 (1.587)	0.147 (0.01)	-0.76 (0.49)
$PhB_{pH6.5}$-$SEDDS_{OliveOil}$	70.18 (3.003)	0.259 (0.009)	-1.27 (0.32)
FaSSIF-SEDDS *Migylol812*	36.41 (1.096)	0.081 (0.014)	-5.35 (0.31)
FaSSIF-$SEDDS_{OliveOil}$	44.76 (0.303)	0.197 (0.002)	-5.73 (0.16)
FaSSIF	62.52 (5.867)	0.234 (0.073)	-14.67 (0.42)

Table 2: Size determination and ζ-potential of the media used in the course of the analysis demonstrating that both SEDDS dispersed uniformly to form stable microemulsions.

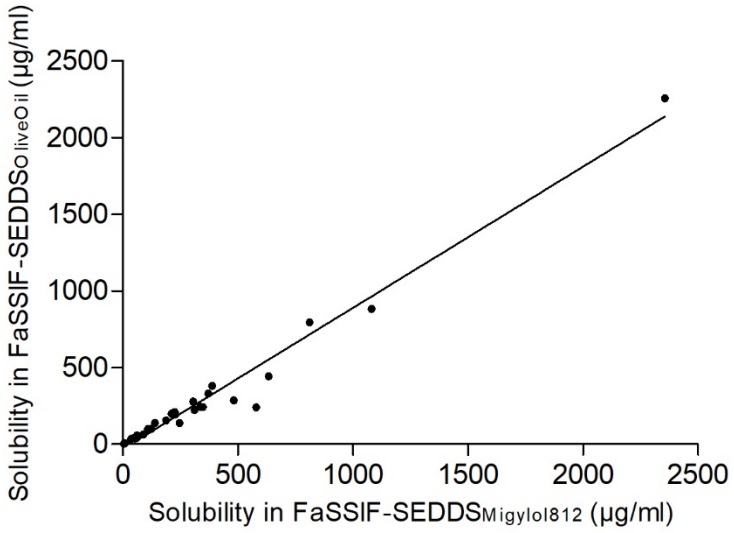


Figure 1*:* Scatter plot of Solubility in FaSSIF-$SEDDS_{Migylol812}$ versus Solubility in FaSSIF-$SEDDS_{OliveOil}$ displaying a high correlation (r^2 = 0.9722). Linear regression line: FaSSIF-$SEDDS_{Migylol812}$ = 42.47 + 1.056(FaSSIF-$SEDDS_{OliveOil}$)

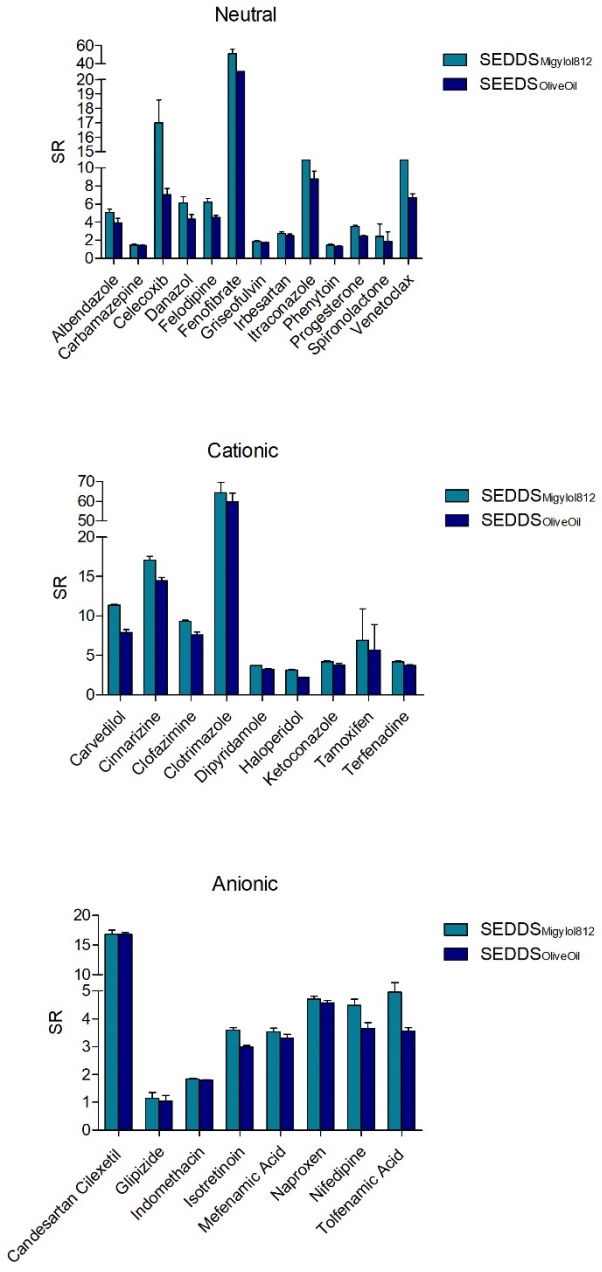


***Figure 2**:* SR (drug solubility in dispersed SEDDS media/FaSSIF) achieved for Neutral, Cationic and Anionic Drugs (pH 6.5). Higher SRs are seen in general for Cationic and Neutral drugs versus Anionic drugs where every anionic drug except Candesartan Cilexetil achieved a SR < 5 in both SEDDS.

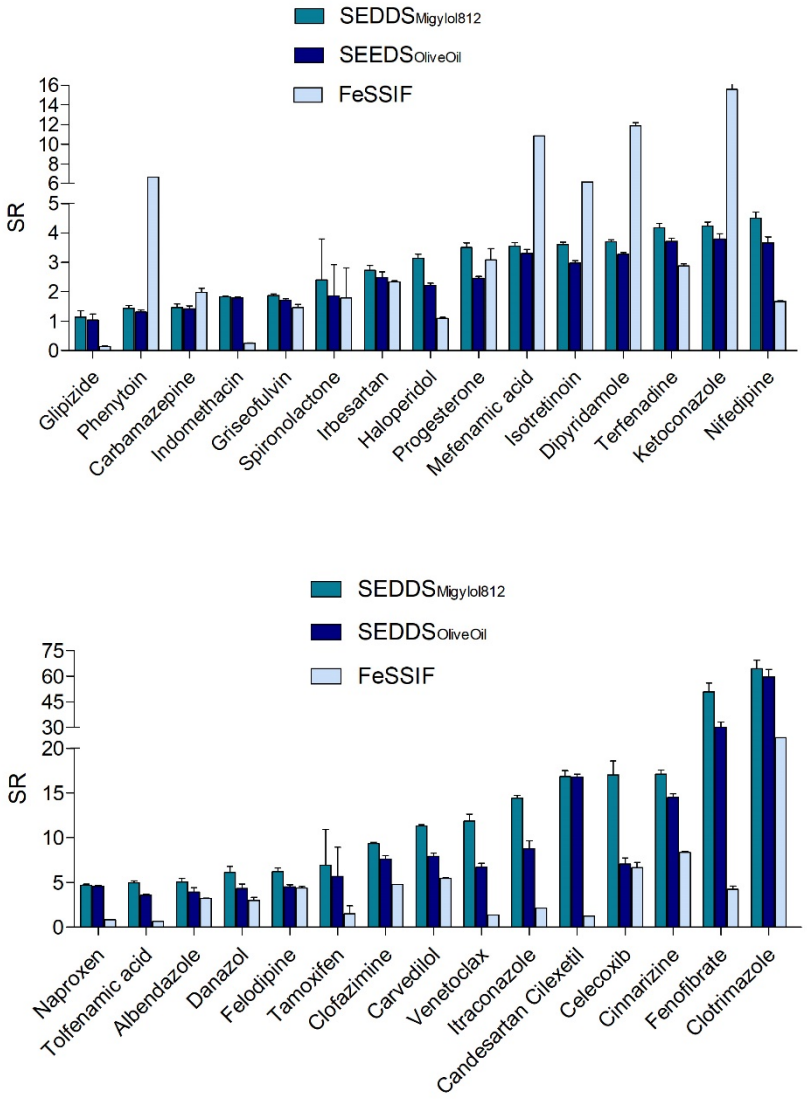


Figure 3: SR (drug solubility in both dispersed SEDDS media/FaSSIF and FeSSIF/FaSSIF). FeSSIF/FaSSIF SR is overcome with $SEDDS_{MC}$ for 24 drugs and with $SEDDS_{LC}$ for 23 drugs, demonstrating ability of the SEDDS to bridge the FeSSIF-FaSSIF solubility gap.

PLS Models

Y-Variable	$logSR_{MC}$	Y-Variable	$logSR_{LC}$
X-Variables	$logD_{6.5}$ MW T_m F_AromB SssCH2	**X- Variables**	$logD_{6.5}$ MW T_m F_AromB SssCH2 NAlip_R
Explained Y-Variance (%)	**81%**	**Explained Y-Variance (%)**	**77%**
No. of PC's	**1**	**No. of PC's**	**2**
RMSEC	**0.19**	**RMSEC**	**0.20**
RMSEP Cross Validation	**0.24**	**RMSEP Cross Validation**	**0.26**
RMSEP Test Set (n=7)	**0.36**	**RMSEP Test Set (n=7)**	**0.37**
r^2 (Calibration)	**0.81**	**r^2 (Calibration)**	**0.77**
r^2 (Validation)	**0.73**	**r^2 (Validation)**	**0.67**

MLR Equations

Y Variable	r^2	$RMSE_{Tr}$	$RMSE_{Te}$ (n=7)	F-value	p-value	Equation
$logSR_{MC}$	0.74	0.23	0.39	16.02	$3.34x10^{-5.}$	$logSR_{MC} = 0.6 + 0.2(logD_{6.5}) + 1.02(F_AromB) - 0.01(T_m)$
$logSR_{LC}$	0.69	0.43	0.37	12.64	$1.37x10^{-4}$	$logSR_{LC} = 0.54 + 0.17(logD_{6.5}) + 1.04(F_AromB) - 0.01(T_m)$

Table 3: **Overview of the PLS models and MLR equations produced for** SR_{MC} and SR_{LC} based on drug descriptors. Tr = Training Set, Te = Test Set. Where calibration refers to the training set and validation refers to the test set. RMSE = Root Mean Square Error. RMSEC = Root Mean Square Error of Calibration. RMSEP = Root Mean Square Error of Prediction.

	FaSSIF	SEDDS$_{Migylol812}$	SEDDS$_{Migylol812}$	SEDDS$_{OliveOil}$	SEDDS$_{OliveOil}$	
Do Equation Used:	**Do**	**Do$_{(SEDDS)}$**	**Do$_{(Predicted)}$**	**Do$_{(SEDDS)}$**	**Do$_{(Predicted)}$**	**DCS Class Transition**
Drug	**DCS Class**	**DCS Class**	**DCS Class**	**DCS Class**	**DCS Class**	
Albendazole	IIb	IIb	IIb	IIb	IIb	
Candesartan Cilexetil	IV	III	IV	III	IV	IV → III
Carbamezapine	IIa	IIa	IIa	IIa	I	
Carvedilol	I	I	I	I	I	
Celecoxib	IIb	I	IIb	IIa	IIb	IIb → IIa/I
Cinnarizine	IIa	I	I	I	I	IIa → I
Clofazimine	IIb	IIa	IIa	IIa	IIa	IIb → IIa
Clotrimazole	IIa	I	I	I	I	IIa → I
Danazol	IIa/IIb	IIa	IIa	IIb	IIa	
Dipyridamole	IV	IV	IV	IV	IV	
Felodipine	I	I	I	I	I	
Fenofibrate	IIb	I	I	I	I	IIb → I
Glipizide	IV	III	III	III	III	IV → III
Griseofulvin	IIb	IIb	IIb	IIb	IIb	
Haloperidol	I	I	I	I	I	
Indomethacin	I	I	I	I	I	
Irbesartan	IV	IV	IV	IV	IV	
Isotretinoin	IIa	I	I	I	I	IIa → I
Itraconazole	IIb	IIb	IIb	IIb	IIb	
Ketoconazole	IIb	IIb	IIa/IIb	IIb	IIb	
Mefenamic Acid	IIb	IIa	IIa/IIb	IIa	IIb	
Naproxen	IIa	I	I	I	I	IIa → I
Nifedipine	IIb	IIa	IIa/IIb	IIa	IIa/IIb	IIb → IIa
Phenytoin	IIb	IIb	IIb	IIb	IIb	
Progesterone	IIb	IIa	IIa	IIa	IIa	IIb → IIa
Spironolactone	IIa/IIb	IIa	IIa	IIa	IIa	
Tamoxifen	I	I	I	I	I	
Terfenadine	IIa	I	I	I	I	IIa → I
Tolfenamic Acid	IIa	IIa	IIa	IIa	IIa	
Venetoclax	IV	III	III	IV	III	IV → III

Table 4: DCS classification of the 30 drugs using both experimental and predicted solubility values. DCS Classes are shown using FaSSIF solubility and both experimentally and predicted solubility's upon SEDDS$_{Migylol812}$ and SEDDS$_{OliveOil}$ dispersion. The different Do equations used to obtain the solubility criteria are also shown.

	FaSSIF	$SEDDS_{Migylol812}$	$SEDDS_{OliveOil}$	
Do Equation Used	**Do**	**$Do_{(Predicted)}$**	**$Do_{(Predicted)}$**	**DCS Class Transition**
Drug	**DCS Class**	**DCS Class**	**DCS Class**	
Clomethiazole Edisilate	IIa	I	I	IIa → I
Dronabinol	IIa	I	I	IIa → I
Ergocalciferol	IIa	I	I	IIa → I
Isotretinoin	IIa	I	I	IIa → I
Cholecalciferol	IIb	I	I	IIb → I
Clofazimine	IIb	IIa	IIa	IIb → IIa
Efavirenz	IIb	I	IIa	IIb → I/IIa
Enzalutamide	IIb	IIa	IIa	IIb → IIa
Ethyl Eicosapentaenoate	IIb	IIa	IIb	IIb → IIa
Fenofibrate	IIb	I	I	IIb → I
Loratidine	IIb	IIa	IIa	IIb → IIa
Menatetrenone	IIb	IIa	IIa	IIb → IIa
Nimodipine	IIb	I	IIa	IIb → I/IIa
Progesterone	IIb	IIa	IIa	IIb → IIa
Teprenone	IIb	IIa	IIa	IIb → IIa
Tocopherol Nicotinate	IIb	I	IIa	IIb → I/IIa
Amprenavir	IV	III	III	IV → III
Nintedanib	IIb	IIb	IIb	
Azithromycin	IV	IV	IV	
Ciprofloxacin	IV	IV	IV	
Cyclosporin A	IV	IV	IV	
Lopinavir	IV	IV	IV	
Ritonavir	IV	IV	IV	
Saquinavir	IV	IV	IV	
Tipranavir	IV	IV	IV	
Vinorelbine Tatrate	IV	IV	IV	

Table 5: DCS Classification of commercial LBF drugs which displayed dose solubility limitation in FaSSIF (Class IIa, IIb, IV) using FaSSIF solubility and both experimentally and predicted solubility's upon $SEDDS_{Migyolol812}$ and $SEDDS_{OliveOil}$ dispersion.

Supplementary Materials

Applying Computational Prediction of Biorelevant Solubility in Dispersions of Self-Emulsifying Lipid-Based Formulations to Predict Dose Number.

Harriet Bennett-Lenane[1], Patrick J. O'Dwyer[1,2], Niklas J. Koehl[1], Karl J. Box[2], Joseph P. O'Shea[1] , Brendan T. Griffin[1]

[1]School of Pharmacy, University College Cork, Cork, Ireland

[2]Pion Inc. (UK) Ltd., Forest Row, East Sussex, UK

*Correspondence to: Brendan Griffin; Tel.: +353 (0) 21 4901657; fax: +353 (0) 21 4901656. Email address: Brendan.griffin@ucc.ie.

1. **Apparent solubility values in $PhB_{Ph6.5,}$ FaSSIF, FeSSIF, FaSSIF-$SEDDS_{Migyolo812}$, FaSSIF-$SEDDS_{OliveOil}$, and SR for $SEDDS_{Migylol812}$, $SEDDS_{OliveOil}$, and FeSSIF versus FaSSIF. FE = Food Effect.**

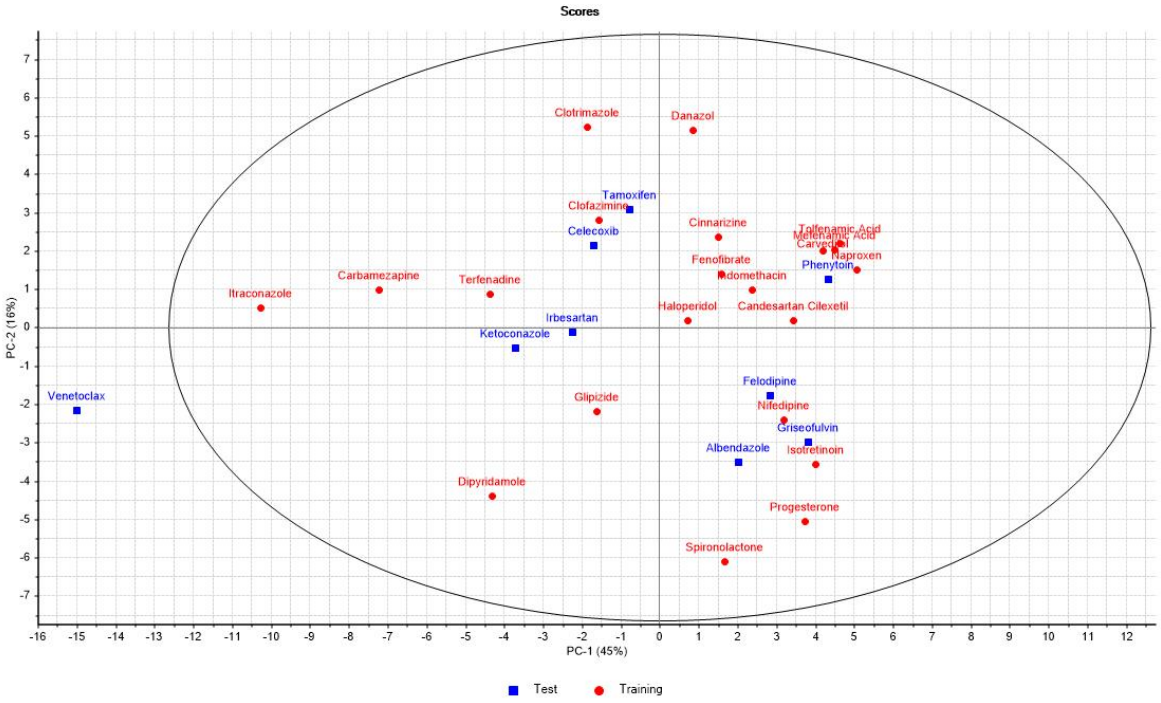
Drug Compound	$PhB_{pH6.5}$ ± SD (µg/mL)	FaSSIF ± SD (µg/mL)	FeSSIF ± SD (µg/mL)	$FaSSIF - SEDDS_{Migylol812}$ ± SD (µg/mL)	$FaSSIF - SEDDS_{OliveOil}$ ± SD (µg/mL)	$logSR_{MC}$	$logSR_{LC}$	logFE
Albendazole	0.9 ± 0.4	1.9 ± 0.0	6.1 ± 0.1	9.62 ± 1.28	7.44 ± 1.634	0.704	0.593	0.507
Candesartan Cilexetil	-	8.26	10	138.9 ± 9.6	138.49 ± 4.63	1.227	1.225	0.083
Carbamazepine	227.1 ± 22.9	266.1 ± 31.4	524.1 ± 25.0	388.127 ± 13.46	379.83 ± 16.47	0.164	0.152	0.294
Carvedilol	46	55.9	305.0 ± 2.0	634.098 ± 5.46	442.88 ± 32.49	1.055	0.898	0.737
Celecoxib	1.54234 ±0.51	34.09 ± 5.12	226	579.82 ± 33.83	240.02 ± 16.07	1.230	0.848	0.823
Cinnarizine	1.4	13.4	112 ± 2.0	228.9 ± 10.84	194.05 ± 9.45	1.232	1.161	0.922
Clofazimine	-	6.2	29.6	57.8 ± 1.68	47.12 ± 4.24	0.969	0.881	0.679
Clotrimazole	2.3 ± 0.3	3.5 ± 0.4	71.1 ± 6.0	225.43 ± 17.8	208.95 ± 11.27	1.809	1.776	1.311
Danazol	0.3 ± 0.05	9.6729 ±1.89	28.8 ± 0.4	59.09 ± 1.44	41.61 ± 3.79	0.786	0.633	0.473
Dipyridamole	6.35	11.56	137.2 ± 6.2	42.90 ± 1.19	37.873 ± 0.907	0.568	0.516	1.074
Felodipine	1.187	54.278	237.0 ± 1.0	337.47 ± 29.1	245.36 ± 14.63	0.794	0.655	0.640
Fenofibrate	0.3 ± 0.0	9.6 ±1.4	40.4 ± 2.9	482.39 ± 47.175	286.55 ± 26.29	1.706	1.480	0.624
Glipizide	22.5 ± 0.6	31.3 ± 3.3	4.3 ± 0.2	35.47 ± 0.62	32.523 ± 1.27	0.054	0.017	-0.862
Griseofulvin	15	20 ±0.9	29.2 ± 3.4	37.35 ± 0.8675	34.15 ± 0.57	0.272	0.233	0.164
Haloperidol	77.81	110.51	120.9 ± 7.3	347.44 ± 25.48	243.65 ± 14.83	0.497	0.343	0.039
Indomethacin	219.0 ± 78.0	443.0 ± 10.0	109.0 ± 7.0	811.48 ± 9.1	794.21 ± 4.77	0.263	0.253	-0.609
Irbesartan	102.0 ± 4.0	112.0 ± 3.4	261	306.29 ± 31.19	277.5 ± 37.22	0.436	0.395	0.367
Isotretinoin	-	52.21	321	188.3 ±6 .90	155.53 ± 6.30	0.556	0.474	0.789
Itraconazole	-	0.33	0.7	4.763 ± 0.16	2.89 ± 0.5	1.159	0.943	0.327
Ketoconazole	6.5	25.91 ± 0.70	403.3 ± 16.5	109.32 ± 5.65	98.21 ± 7.43	0.625	0.579	1.192
Mefenamic acid	-	60	649	212.46 ± 13.292	198.6 ± 13.19	0.549	0.520	1.034
Naproxen	230.26	492.29	401	2356.17 ± 95.78	2255.05 ± 75.76	0.672	0.661	-0.089
Nifedipine	11.5	27.8	46.1 ± 1.0	124.78 ± 9.93	101.72 ± 9.65	0.652	0.564	0.220
Phenytoin	39.07	42.84	283	61.51 ± 7.24	56.29 ± 5.58	0.158	0.117	0.820
Progesterone	11.16	25.56	78.6 ± 16.2	89.34 ± 5.98	62.72 ± 1.21	0.544	0.389	0.488
Spironolactone	22	25.8	46.0 ± 2.5	61.9 ±7.08	47.8 ± 6.05	0.380	0.267	0.251
Tamoxifen	5.9	156	236.0 ± 13.0	1081.47 ± 56.36	882.21 ± 20.79	0.839	0.753	0.180
Terfenadine	13.6 ± 1.3	89.0 ± 4.0	256	371.23 ± 15.46	329.95 ± 4.28	0.620	0.569	0.459
Tolfenamic acid	27.404	62.779	41.0 ± 0.5	311.56 ± 16.03	224.37 ± 7.13	0.696	0.553	-0.185
Venetoclax	0.04	20.729 ±0.51	28.4 ± 2.2	246.340 ± 25.75	138.83 ± 16.37	1.075	0.825	0.137

2. RP-HPLC/UV methods for the 6 drugs completed using the Shake Flask Method with HPLC-UV analysis.

Drug	Column	A	B	Ratio	Temp (°C)	Flow Rate (ml/min)	Inj. Vol (µL)	λ (nm)
Danazol	Symmetry C18 5 µm, 4,6 x 150 mm	ACN	Water	55:45	25	1	50	286
Ketoconazole	Symmetry C18 5 µm, 4,6 x 150 mm	Phosphate buffer 10 mM, pH 8.5	ACN	40:60	25	0.8	50	297
Venetoclax	Zorbax Eclipse Plus-C18 column (5 µm, 4.6 mm x 150 mm) including Zorbax 156 Eclipse Plus-C18 guard column (5 µm, 4.6 mm x 12.5 mm)	ACN + 0.5 % TFA	Water + 0.5 % TFA	53:47	40	1	50	316
Fenofibrate	Symmetry C18 5 µm, 4,6 x 150 mm	NaAc 25 mM, pH 5.0	ACN	20:80	25	1	50	287
Celecoxib	Symmetry C18 5 µm, 4,6 x 150 mm	ACN + 0,15%TEA, pH3	Water	55:45	25	1	20	254
Griseofulvin	Symmetry C18 5 µm, 4,6 x 150 mm	ACN	Water	55:45	25	1	50	292

3. **Principle Component Analysis (PCA) scores plot detailing the chemical space occupied by the Training and Test Sets of the dataset.**

61% of the variation in the dataset is explained by PC-1 and PC-2 and Venetoclax is outside the 95% confidence level and was therefore placed in the test set.



4. **Preliminary Studies Testing the Two Solubility Methods Used.**

Drug Solubility in FaSSIF-SEDDS$_{Migylol812}$ and FaSSIF-SEDDS$_{OliveOil}$ completed for both shake flask and µDISS methods using Danazol. Solubility's were obtained using FaSSIF-V2 for the µDISS method, which contains a smaller concentration of lecithin, due to powder availability at that time, therefore a ratio of solubility in MC/LC was calculated to test similarity of results instead of direct comparisons.

Results:

	Shake Flask	**uDiss**	**Ratio MC/LC Solubility**
FaSSIF-SEDDS$_{Migylol812}$	59.089 µg/mL (±1.44)	37.130 µg/mL (±0.589)	Shake Flask = 1.42
FaSSIF-SEDDS$_{OliveOil}$	41.612 µg/mL (± 3.79).	22.878 µg/mL (±1.138)	µDISS = 1.6