

OPEN

dendPoint: a web resource for dendrimer pharmacokinetics investigation and prediction

Lisa M. Kaminskas^{1,6*}, Douglas E. V. Pires^{2,3,4,6} & David B. Ascher^{3,4,5*}

Nanomedicine development currently suffers from a lack of efficient tools to predict pharmacokinetic behavior without relying upon testing in large numbers of animals, impacting success rates and development costs. This work presents dendPoint, the first *in silico* model to predict the intravenous pharmacokinetics of dendrimers, a commonly explored drug vector, based on physicochemical properties. We have manually curated the largest relational database of dendrimer pharmacokinetic parameters and their structural/physicochemical properties. This was used to develop a machine learning-based model capable of accurately predicting pharmacokinetic parameters, including half-life, clearance, volume of distribution and dose recovered in the liver and urine. dendPoint successfully predicts dendrimer pharmacokinetic properties, achieving correlations of up to $r = 0.83$ and Q^2 up to 0.68. dendPoint is freely available as a user-friendly web-service and database at <http://biosig.unimelb.edu.au/dendpoint>. This platform is ultimately expected to be used to guide dendrimer construct design and refinement prior to embarking on more time consuming and expensive *in vivo* testing.

A lack of appropriate pharmacokinetic behavior has historically been one of the leading causes of drug failure in clinical trials¹. Advances in controlled release technologies and nanomedicine, however, are increasingly providing new opportunities to circumvent this shortcoming in rational drug development initiatives and are providing renewed hope for old drug candidates. Importantly, a wide range of nanosized materials with a variety of chemico-biological traits that can be used to alter and drive the pharmaceutical behavior of loaded drugs (including colloids, nanoparticles and polymers) have been developed and explored for their potential as relatively biologically inert drug carriers. One of the primary indications for which nanomaterials have been explored and have proven successful is in promoting the targeted delivery of chemotherapeutic drugs towards solid tumors via the enhanced permeation and retention (EPR) effect². While optimal EPR necessitates the use of nanocarriers that display prolonged blood circulation³, a trade off needs to exist between blood exposure (to maximize EPR) and elimination (to minimize accumulation of the nanomaterial in the body and off-target toxicity). As a classic example, the PEGylated liposomal formulation of doxorubicin (Doxil[®]/Caelyx[®]) displays good EPR into solid tumors, but its prolonged plasma exposure leads to accumulation in the extremities, causing painful swelling of the hands and feet⁴. This highlights the importance of optimizing the pharmacokinetic behavior of nanomedicines early in development.

Often, however, the preclinical development of nanomedicines involves testing the biopharmaceutical behavior and safety of a wide range of nanocarriers, alone and in combination with loaded drug, in hundreds of animals prior to advancing the optimized construct(s) into clinical trials. This leads to increased research and development time and cost which, ultimately, translates into higher product costs for consumers. Furthermore, researchers are seeing an increasing impetus to limit the use of animals in biomedical research⁵. This has been addressed somewhat for small molecule drug candidates by the development of predictive models for toxico-pharmacokinetic behavior based on the physicochemical attributes of the drug (approaches such as pkCSM⁶). This enables preliminary *in silico* assessment of pharmacokinetic properties, guiding refinement of the molecule prior to *in vivo* testing. To date, however, no such predictive models exist for macromolecules and

¹School of Biomedical Sciences, University of Queensland, St Lucia, Queensland, Australia. ²School of Computing and Information Systems, University of Melbourne, Melbourne, Victoria, Australia. ³Computational Biology and Clinical Informatics, Baker Heart and Diabetes Institute, Melbourne, Victoria, Australia. ⁴Structural Biology and Bioinformatics, Department of Biochemistry, Bio21 Institute, University of Melbourne, Melbourne, Victoria, Australia. ⁵Department of Biochemistry, University of Cambridge, Cambridge, Cambridgeshire, United Kingdom. ⁶These authors contributed equally: Lisa M. Kaminskas and Douglas E. V. Pires. *email: l.kaminskas@uq.edu.au; david.ascher@unimelb.edu.au

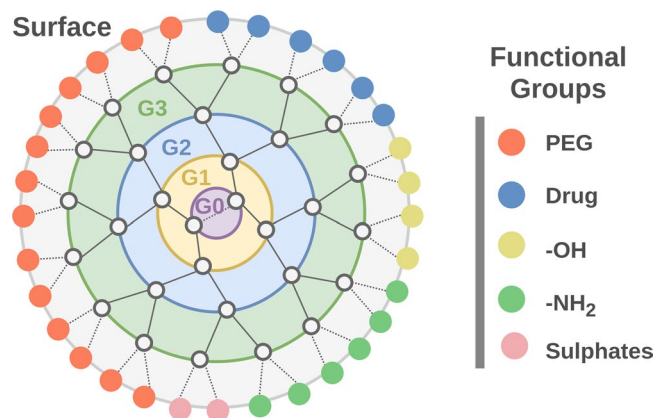


Figure 1. Basic structure of a dendrimer showing sequential layering of monomeric units around a central core (G0). A dendrimer may be composed of any monomeric unit provided it has at least 2 functional groups available to build additional generations. Surface functional groups depicted as circles.

nanomaterials. This is in part due to the wide diversity in available nanostructures that can be employed as drug delivery systems, with each displaying distinct *in vivo* behavior. Even within defined classes of nanomaterials, changes to the nanomaterial composition, drug loading, length and number of surface polyethylene glycol (PEG) groups, for instance, can have profound and, until recently, seemingly unpredictable effects on biopharmaceutical behavior by altering the solution behavior and cell/protein binding properties of the material⁷. This is especially problematic for polymer-based systems (linear and hyperbranched polymers) which are typically much smaller (≤ 20 nm or < 500 kDa) than colloids and nanoparticles (typically > 100 nm) and are therefore, more sensitive to small changes in composition and physicochemical properties.

In an attempt to address the lack of effective predictive models for the *in vivo* behavior of nanomaterials, Riviere and colleagues⁸ published the first approach to predict the adsorption of biomolecules onto a nanoparticle surface in Nature in 2010. The approach involved comparing the surface adsorption of a set of small molecule probes and generating a ‘surface adsorption index’ to predict the binding of biomolecules (the ‘protein corona’) which is known to play a significant role in dictating the biodistribution behavior of nanoparticles⁹. Subsequent to this, a number of investigators have used physiologically based pharmacokinetic models (PBPK) to simulate the mass-time bio-distribution profiles for a range of metal nanoparticles^{10–15} as well as some polymeric nanoparticles^{16–18}. In most cases, these models were developed based on limited experimental data sets to predict the biodistribution and elimination kinetics of nanoparticles with a fairly narrow set of physicochemical variants (such as size and charge). The intention behind these models was to aid researchers in their selection of optimal particle properties for further development or in risk assessment analysis. The PBPK approach however, is not appropriate for predicting the pharmacokinetic behavior of more complex nanostructures such as liposomes and polymers that may be comprised of a variety of different scaffold components (such as different lipids or monomers). These models are also not easily adaptable and available for use by researchers with limited or no knowledge of biometric analysis.

Dendrimers are well defined hyperbranched polymeric systems that can range in size from 1–20 nm in diameter¹⁹ (Fig. 1), which can provide several pharmacokinetic advantages over much larger colloids and nanoparticles^{20–22}. Drugs can be loaded either peripherally via internally triggered chemical linkers, or can be non-covalently loaded into the hydrophobic scaffold. Although the clinical advancement of nanomedicines has been a slow process, Starpharma’s topical microbicidal gel (Vivagel[®]) has recently gained regulatory approval in Australia and Europe for the treatment of bacterial vaginosis and a dendrimer-based formulation of docetaxel (DEP[™]-docetaxel) recently successfully completed phase I clinical trials for the treatment of advanced solid tumors. The establishment of an *in silico* model capable of accurately predicting dendrimer pharmacokinetics is therefore timely and of increasing relevance.

Here, we describe dendPoint, the first *in silico* and widely available model to predict the intravenous pharmacokinetics of complex polymeric nanomaterials based on scaffold structure and physicochemical properties. We have manually curated a detailed relational database describing dendrimer biopharmaceutical behavior with various structural and chemical characteristics. This was used to develop a model to predict key pharmacokinetic parameters for dendrimers. dendPoint is available via a user-friendly freely available web-based system, accessible at <http://biosig.unimelb.edu.au/dendpoint>. This computational platform encompasses a relational database of pharmacokinetic properties of different dendrimer scaffolds together with a web-service capable of predicting and comparing dendrimer properties, including Half-life, Volume of Distribution, Clearance and Dose in Liver and Urine, allows users to rapidly and easily browse literature-derived properties as well as predict, compare and visualize dendrimer pharmacokinetic properties.

Results

Database curation. In total, the pharmacokinetic parameters of 69 distinct dendrimers, from over 600 papers, were manually curated into the dendPoint database (Fig. 2, Table S1). Many structural and physicochemical properties can dictate the pharmacokinetic behavior of dendrimers, including scaffold composition, dendrimer size, degree of surface PEGylation, PEG chain length, surface functionality (including charge and the

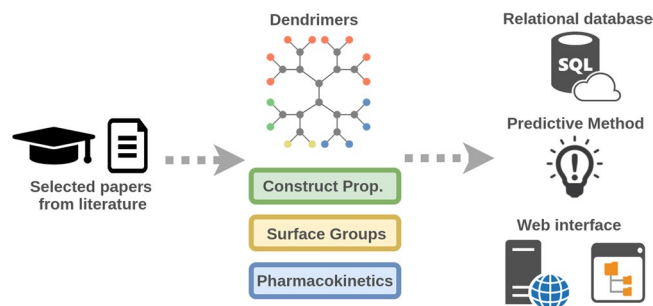


Figure 2. dendPoint workflow. Dendrimer pharmacokinetics were collected from over 600 papers via literature search. Construct properties, information on surface functional groups as well as pharmacokinetic behaviour of 69 different dendrimers were collected and included in a relational database. This was used as evidence to train and test predicted methods via supervised learning. A user-friendly web interface was created for both database and predictive method.

presence of hydrophobic drugs) and structural flexibility^{3,23,24}. The impact of each of these parameters on intravenous pharmacokinetics has been summarized previously³ and they were therefore included in the database as summarized in Table S1 (Supporting Information).

The susceptibility of the dendrimer scaffold to *in vivo* biodegradation may also impact upon the rate of dendrimer elimination from the body, but surface functionalization with non-biologic groups (such as non-natural amino acids and PEG) slows scaffold breakdown³. To this end, with the exception of two unmodified amine-terminated polylysine dendrimers, none of the dendrimers in Table S1 were reported in their respective publications to have shown significant *in vivo* biodegradation that was expected to have driven the reported intravenous pharmacokinetics. Biodegradability of the scaffold was, therefore, not included as a parameter that defined the ultimate pharmacokinetics, although the composition of the scaffold was.

Pharmacokinetic data was therefore obtained for dendrimers based on non-biodegradable triazine and poly-amidoamine (PAMAM) scaffolds, as well as biodegradable polyester and polylysine scaffolds. While the effective ‘size’ of polymeric nanomaterials may be reported in terms of hydrodynamic radius or molecular weight, few of the papers reported in Table S1 provided this information on radius and therefore, this parameter could not be included in the database. Regardless, it has previously been suggested that the terminal Half-life of dendrimers correlates more significantly with molecular weight than with radius³. In addition, three quarters of the dendrimers included in dendPoint presented some degree of surface PEGylation. This is likely a result of the fact that since the molecular weight of the dendrimer scaffold is limited by poor conjugation efficiency and greater polydispersity for generations higher than approximately 5–6, surface PEGylation is commonly required to increase size and prolong plasma exposure³. PEGylation is also often employed either alone or in combination with acetylation (Ac) to block surface reactive sites and prevent binding to cells and tissues.

Pharmacokinetic parameters that were included in the database include terminal Half-life and clearance, since these ultimately describe the plasma exposure of the dendrimers after intravenous delivery. From these parameters, volume of distribution can be calculated from the equation $Cl = Vd^* (0.693/t_{1/2})$, where Cl denotes Clearance, Vd denotes terminal Volume of Distribution and $t_{1/2}$ denotes Half-life. The percentage of the dose excreted via the urine and percentage dose recovered in liver were also included since these represent the major pathways by which dendrimers are cleared from plasma (*i.e.*, via elimination in urine and biodistribution towards the liver). For liver uptake and urinary excretion, a threshold of <20% was used in classification tasks to define dendrimers with limited liver uptake or urinary excretion respectively. With the exception of several polyester dendrimers, significant quantities of dendrimer have not been detected in the feces, suggesting this is not a major route of dendrimer elimination from the body. Liver biodistribution typically results from recognition of the polymeric nanomaterial by macrophages of reticuloendothelial organs (which also include the spleen, lymph nodes and lungs) and is normally the organ that contains the highest proportion of an injected dendrimer dose after one week³. Distribution towards the liver may result from initial plasma protein binding (opsonisation) of the dendrimer, electrostatic recognition of anionic charges on the dendrimer surface by macrophages or non-specific accumulation of long circulating constructs over time^{23–25}.

Analysis of dendrimer properties. The distribution of all four experimental pharmacokinetic properties for all molecules in the database are shown in Fig. S1. This highlights the relatively broad distribution of Half-life, Clearance and percentage dose recovered in the urine across the range of dendrimer constructs that have been characterized. There was little accumulation of dendrimer dose (median of 7%) in the liver for the majority of constructs in the database (with 80% of dendrimer having a %Dose in Liver below 20%).

Looking closer at the distribution of Half-life and Clearance divided by Scaffold and Flexibility revealed some general trends (Fig. 3A–D). Notably, as structural flexibility increased, Half-life decreased and Clearance increased. When assessing the effects of Surface Charge on the pharmacokinetic properties (Fig. 3E,F) a similar behavior was also observed. While Clearance increased as Surface charge move further from neutral (increasing charges, either negative or positive), Half-life decreased. The polylysine-based scaffold presented the largest variability in pharmacokinetic properties, in part due to the large number of diverse constructs that have been systematically analyzed to date. Triazine-based dendrimers were associated with lower Clearance and longer Half-lives.

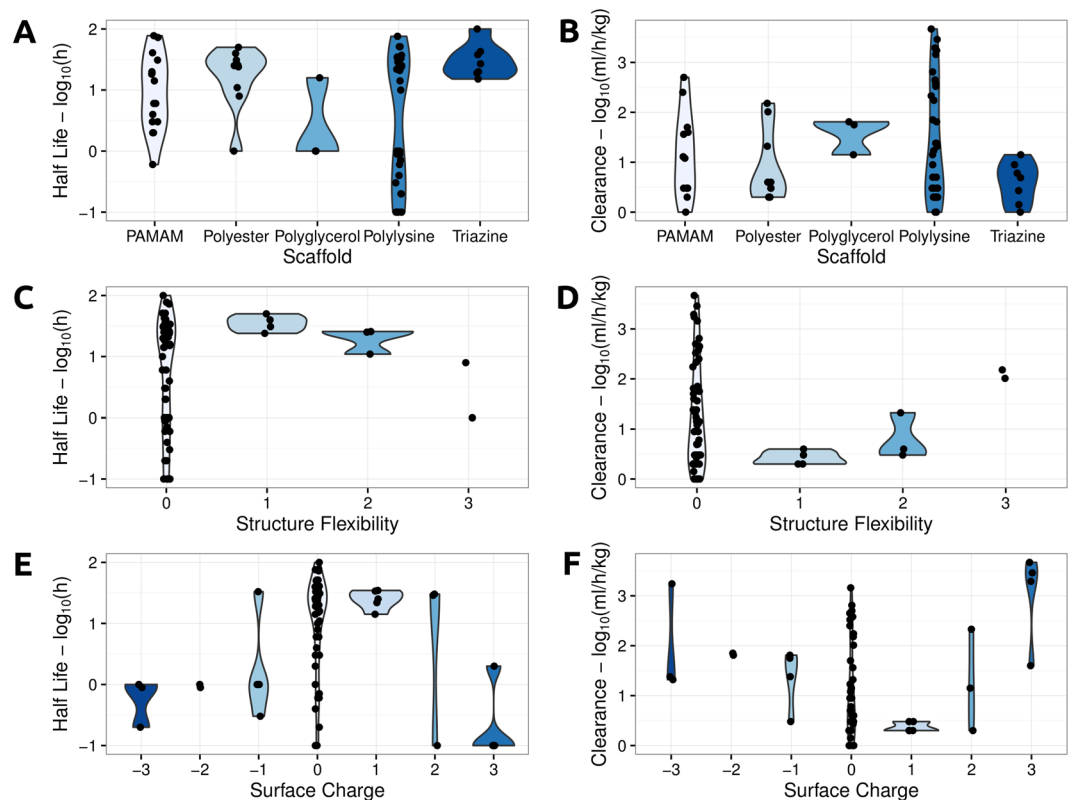


Figure 3. Distribution of Half-life and Clearance properties based on Scaffold, Structure Flexibility and Surface Charge across the database. The left-hand side graphs depict (from top to bottom), as violin plots, the distribution of Half-life per Scaffold, Structure Flexibility and Surface Charge, while the right-hand side graphs show how Clearance varies based on these properties.

PEGylation level and construct molecular weight correlated well with all pharmacokinetic properties. Half-life significantly and positively correlated with both construct molecular weight ($r = 0.70$) and total PEG molecular weight ($r = 0.61$) (Fig. 4A,B). An even stronger correlation, although negative, between these properties and Clearance was also observed ($r = -0.70$ for construct molecular weight and $r = -0.65$ for total PEG molecular weight) (Fig. 4C,D).

Prediction of dendrimer pharmacokinetic properties. Our curated database was then used to train and test predictive pharmacokinetic models. The approach we used was based on our well validated CSM methodology²⁶. This is based on the concept of structural signatures, which are an alternative way of extracting relevant patterns from molecular entities, originally modeled as graphs, which in turn are provided as evidence to supervised learning methods. These structural signatures are a powerful and scalable way to represent geometry and physicochemical properties, and have been applied to accurately predict small molecule pharmacokinetics^{6,27}, to characterize small molecule-protein interactions²⁸ and the effects of mutations on protein structure²⁹⁻³⁷. As shown in Fig. 1, we represent each dendrimer as a graph where the nodes are the branch points and the edges are the connections. Distance patterns between nodes are then summarised as cumulative distribution functions, which are then used as evidence to train machine learning methods. Complementary information also integrated into the signatures included the physicochemical properties in the curated database. This information, together with the experimentally measured pharmacokinetic properties, was then used to train and test predictive models.

The dendPoint platform for predicting dendrimer pharmacokinetic properties was able to accurately predict Half-life with a Pearson's correlation coefficient of $r = 0.82$ and $Q^2 = 0.66$ on jackknife validation (Table 1, Fig. 5A). This correlation increases to $r = 0.91$ when assessing the performance of the method after removing 10% of outliers. A similar performance, however with smaller dispersion, was observed for the Clearance predictor. A correlation of $r = 0.83$ and $Q^2 = 0.68$ was obtained on cross validation, also increasing to $r = 0.89$ after 10% outlier removal (Table 1, Fig. 5B). To further evaluate the predictive models, we assessed their performance on a bootstrap validation using a 90%/10% split over 100 repetitions. The performance of all models was consistent with that achieved over jackknife validation as shown in Table 1.

For both Half-life and Clearance predictors, the outlier sets were composed, on average, of smaller constructs (Construct molecular weight of Half-life outliers = 33.8 kDa, and Clearance outliers = 17.2 kDa, compared to the average construct molecular weight across the dataset of 47.9 kDa), with less surface PEGs (Half-life outliers = 8.5, Clearance outliers = 6.4 compared to 16.2 average surface PEGs across the dataset), accounting for a

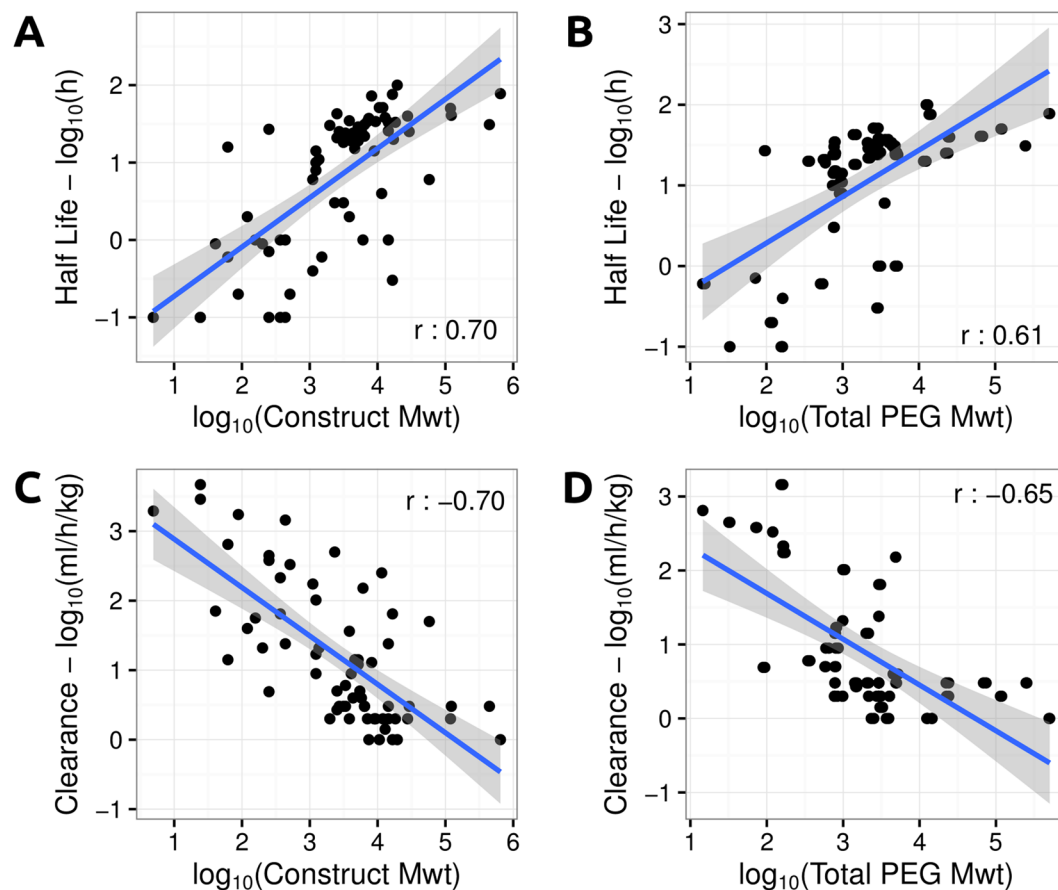


Figure 4. Distribution of Half-life and Clearance properties based on Construct and Total PEG molecular weight. The left-hand side graphs depict, as scatter plots, the distribution of Half-life (top) and Clearance (bottom) based on Construct molecular weight, while the right-hand side graphs show how these properties vary based on Total PEG molecular weight. Pearson's linear correlations between these properties are also depicted.

| Data set | Jackknife Cross Validation | | Bootstrap | |
|-------------|----------------------------|------|-----------|------|
| | Pearson's | RMSE | Pearson's | RMSE |
| Half-life | 0.82 | 0.52 | 0.76 | 0.54 |
| Clearance | 0.83 | 0.57 | 0.82 | 0.54 |
| %Dose Liver | 0.59 | 15.7 | 0.57 | 13.6 |
| %Dose Urine | 0.73 | 15.5 | 0.70 | 14.2 |

Table 1. dendPoint evaluation on jackknife cross validation, and 100 times bootstrap validation on a 90%/10% split with 100 repetitions.

smaller total PEG molecular weight (Half-life outliers = 15.9 kDa, Clearance outliers = 11.5 kDa compared to the overall dataset that had an average total PEG molecular weight of 30.3 kDa).

Building a predictor for %Dose in liver presented a great challenge given the very skewed distribution of experimental values, as seen in Fig. S1C. Despite the skewed distribution of experimental data, dendPoint was able to achieve a correlation of $r = 0.59$, which increased significantly once the top 10% of outliers were removed, reaching a correlation of $r = 0.73$ on 90% of the data (Table 1, Fig. S2A). Predictive performance for %Dose Liver did deteriorate slightly for larger predicted values due to the skewed nature of the data distribution.

The predictor for %Dose in urine, achieved a higher correlation in comparison with % Dose in Liver, with a correlation of $r = 0.73$, increasing to $r = 0.87$ after 10% outlier removal (Table 1, Fig. S2B). This is largely due to the data of %Dose recovered in the urine within the dendPoint database that had a more even distribution of experimental values. Bootstrap validation of both %Dose in liver and %Dose in urine showed consistent performance to jackknife (Table 1), further improving confidence in both models predictive capabilities.

Alternatively, we were also able to build predictors to assess whether a dendrimer construct would be cleared by the liver and/or excreted in urine. Constructs were defined as having limited liver uptake or urinary excretion using a 20% cutoff for dose in liver and urine, respectively (see details in Materials and Methods). Figure S3 of

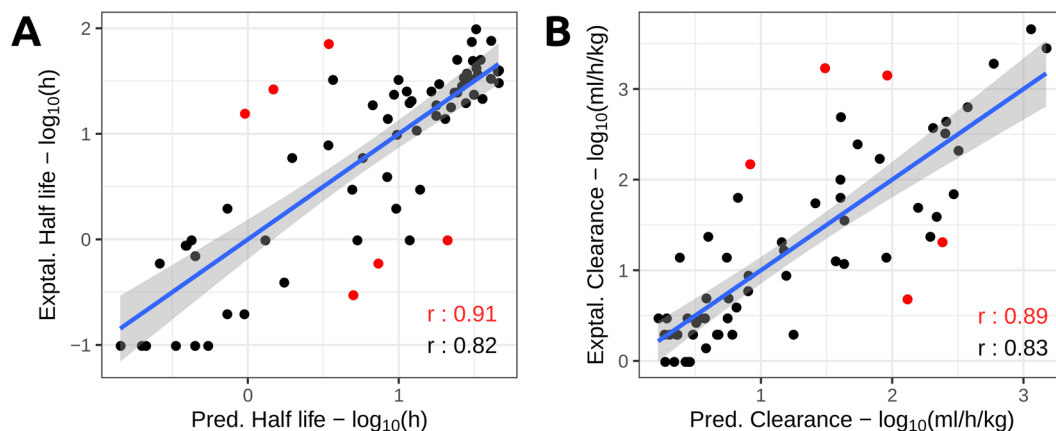


Figure 5. Predicting dendrimer pharmacokinetics with dendPoint. The graphs depict the correlation between experimental and predicted properties on jackknife validation. A correlation of $r = 0.82$ was obtained when assessing Half-life, increasing to $r = 0.91$ when 10% outliers (shown in red) were removed. While predicting Clearance, dendPoint achieves a correlation of up to $r = 0.89$ on 90% of the data.

Supplementary Materials shows the ROC curves obtained for both predictors, which achieved AUCs of 0.87 and 0.86 for urinary excretion and liver uptake, respectively. The predictors were successful in predicting the differences between constructs that had limited liver uptake/limited urinary excretion compared to those that were cleared by the liver or excreted in urine, achieving an accuracy of up to 80%.

The attributes that contributed most to the performance of each predictor were evaluated by Principal Component Analysis (PCA). Both construct molecular weight and total PEG molecular weight were consistently well ranked for all four predictors. This was expected given these attributes correlated well with the pharmacokinetic parameters by themselves. PCA showed that these attributes contributed largely to the variability of the Half-life data set, together with Generation and Structure Flexibility (Fig. S4A). These were consistent with the other data sets. Drug conjugation, both in terms of the type and number of surface drugs, played an important role in prediction performance, despite a clear correlation not being noticed during analysis of individual features. A histogram of the percentage of explained variance per feature (Fig. S4B) shows that the majority of the features are necessary to explain variability (a linear drop on explained variance, instead of a usual logarithmic drop), suggesting that the selected group of variables are diverse and complementary.

Discussion

In summary, here we describe dendPoint, the first relational database and predictive method that associates physicochemical properties of a complex hyperbranched polymeric structure (notably dendrimers) with experimentally measured intravenous pharmacokinetic data. This provides the first opportunity to begin to systematically analyze the relationship between dendrimer structures and their biological behaviors, in the attempt to guide construct design and development. It has been carefully curated from the literature and will be updated regularly. Although, in practice, dendrimers may ultimately be delivered via non-intravenous routes (such as subcutaneously or via inhalation) which will require the need for additional base physicochemical properties for optimal pharmacokinetic behavior, the intravenous route is currently standard practice for the systemic delivery of nanomedicines.

This database reveals some general rules of dendrimer design, where construct molecular weight, flexibility and PEGylation can all be used to tunably adjust the plasma exposure of a dendrimer. To begin with, whilst half-life is used as the most common parameter to describe the plasma exposure of a nanomaterial, clearance is a more appropriate parameter since it takes into account urea under the whole plasma concentration-time profile, rather than simply the elimination kinetics alone. With this in mind, plasma exposure can be increased (or rather, clearance decreased) via restricting urinary excretion, extravasation and uptake into cells and tissues by employing the following basic rules: (1) increased degree of surface conjugation or molecular weight of hydrophilic, biocompatible and poorly-biodegradable polymers such as PEG (polymerisation), (2) increase construct molecular weight (size), (3) reduce surface charge (charge), (4) reduce structural flexibility (flexibility). Based on the available data and our model, drug conjugation to the surface has a negligible effect on intravenous pharmacokinetics.

In small molecule drug design, the development and application of generalized rules and these tools have been widely used to improve compound quality and success rates. In addition to providing the first curated database of dendrimers to facilitate the analysis of nanoparticles, we demonstrate that it can be used as the basis to train novel predictive pharmacokinetic models.

We have implemented a user-friendly web server that will enable researchers to search, predict and visualize the pharmacokinetic properties for their molecules of interest (<http://biosig.unimelb.edu.au/dendpoint>). In addition, we have implemented a comparison feature that enables users to rapidly compare the pharmacokinetic profiles of two molecules, allowing systematic evaluation of pharmacokinetic profiles as various physicochemical properties of the dendrimer are modified. Considering the sensitive nature of many projects, the web server does not retain any information submitted to it. This will hopefully facilitate the development and optimization of dendrimers for specific biological roles, and provide a foundation for the evaluation of nanoparticles more broadly.

Methods

Collation of published pharmacokinetic data. Published work describing the intravenous pharmacokinetics of dendrimers were identified by undertaking PubMed searches of the terms ‘dendrimer and pharmacokinetics’ or ‘dendrimer and biodistribution’. Only papers that described - or provided enough information to extrapolate - composition of the dendrimer scaffold and surface, molecular weight, terminal plasma half-life ($t_{1/2}$) plus % dose excreted in urine and/or % dose recovered in liver at termination were included in the database. Papers describing the pharmacokinetics of loaded drug rather than the dendrimer scaffold were excluded, since following the loaded drug is a poor predictor of the biopharmaceutical behavior of the dendrimer itself. Supramolecular structures (such as micelles) that were comprised of dendritic polymers and hyperbranched polymers that were based on dendritic cores were also excluded. Dendrimers conjugated with distinct targeting moieties (such as tumor targeting ligands) were also not included since this is expected to change the pharmacokinetics of the base construct.

Using this criteria, 20 papers collectively describing the pharmacokinetics of 69 distinct dendrimer structures were identified that provided sufficient information to compile the database. Plasma clearance (Cl) was included in the database where possible and was either taken directly from published papers or was extrapolated from available data by dividing dose by area under the plasma concentration versus time curve. Where pharmacokinetic parameters needed to be extrapolated from graphical data, the data recovery program GetData Graph Digitizer v2.26 (GetData Pty Ltd, NSW, Australia) was employed. Where area under the plasma concentration time curve needed to be calculated to determine Cl , this was calculated manually using the trapezoid rule and extrapolated to infinity by dividing the last plasma concentration detected by the elimination rate constant (k).

Each of the papers used described dendrimer pharmacokinetics in rodent models (notably mice and rats), and as such, the body weight of the animal models used can vary by up to 20 fold. While this has no bearing on $t_{1/2}$ or proportion of injected dose recovered in liver or urine, Cl is a function of distribution volume (Vd) and therefore body weight. Cl was therefore normalized to ml/h per kg body weight by dividing Cl (in ml/h) by the average reported body weight of animals used in the study. Where mean body weight or body weight ranges were not reported, mean body weights for the animals described and at the reported ages were extrapolated from growth curves published online by several breeders (Harlan, Charles River Laboratories, Taconic and the Animal Resources Centre).

The complete set of data used to compile the pharmacokinetic database is tabulated in Table S1 (see Supporting Information).

Method training and evaluation. Different supervised learning algorithms for regression available on the Weka ToolKit (version 3.8.2)³⁸ were used to train models for the different pharmacokinetic properties collected and store on the database. The best performing models, based on the evaluation metrics below, were obtained using Random Forest³⁹ (default settings, 100 trees) for both classification and regression tasks.

Regression models were initially evaluated using jackknife validation procedure and the best performing models were selected based on the Pearson's Correlation Coefficient and Root Mean Square Error (RMSE);

$RMSE = \sqrt{\frac{\sum_{i=1}^N (Predicted_i - Actual_i)^2}{N}}$. The Pearson's Correlation Coefficient quantifies the linear dependency between two variables (e.g., experimental vs. predicted pharmacokinetic properties) as the covariance of the variables divided by the product of their standard deviations: $r_{A,B} = cov(A,B)/sd(A)*sd(B)$. Correlations vary from $[-1, 1]$, where 1 and -1 denote a perfect positive and negative linear correlation, respectively, while a correlation value of 0 denote no linear correlation. Classification models were also evaluated using jackknife validation and best performing ones selected based on Area Under ROC curve (AUC) and accuracy. AUC varies from $[0, 1]$ with a random binary classifier achieving an AUC = 0.5 and a perfect classifier achieving an AUC = 1. Regression and classification models were further evaluated using bootstrap validation on a 90%/10% split of the data, with 100 repetitions. Outliers are considered the points furthest away from the line of best fit, and were removed only for analysis purposes. Accuracy is denoted by the proportion of correctly classified instances. PCA was performed using R programming language to assess each feature's contribution to explain variability.

Database and web interface. dendPoint's database and predictive models have been implemented as a user-friendly web-server freely available at <http://biosig.unimelb.edu.au/dendpoint>. The dendrimer pharmacokinetics information collected from literature search was consolidated as a MySQL relational database (version 5.5.35). Front-end development was created using the Bootstrap framework (version 3.3.7). The server back-end runs on a Linux server and was implemented in Python using the Flask framework (version 0.12.3). Dendrimer depiction was developed with the JavaScript library D3.js (version 4.2.6) and Plasma Concentration plot was created with the HighCharts charting tool (version 5.0.0).

Users have the option to browse the database via the web interface (Fig. S5, Supporting Information), search/filter specific information as well as show/hide construct properties, surface group properties and pharmacokinetics details.

Job submission can be easily done by informing construct properties and surface group compositions via an intuitive submission form (Fig. S6). Generation and Construct molecular weight are required fields. After prediction a results page is exhibited (Fig. S7), showing the pharmacokinetics properties, dendrimer depiction and plasma concentration curve, giving the user the option to wither modify the job for resubmission or compare the predicted properties with another dendrimer construct (Fig. S8). See Supporting Information for supplementary figures.

Received: 14 September 2018; Accepted: 24 September 2019;

Published online: 29 October 2019

References

- Prentis, R. A., Lis, Y. & Walker, S. R. Pharmaceutical innovation by the seven UK-owned pharmaceutical companies (1964–1985). *British journal of clinical pharmacology* **25**, 387–396 (1988).
- Svenson, S. Clinical translation of nanomedicines. *Curr Opin Solid St M* **16**, 287–294, <https://doi.org/10.1016/j.cossms.2012.10.001> (2012).
- Kaminskas, L. M., Boyd, B. J. & Porter, C. J. Dendrimer pharmacokinetics: the effect of size, structure and surface characteristics on ADME properties. *Nanomedicine* **6**, 1063–1084, <https://doi.org/10.2217/nnm.11.67> (2011).
- Charrois, G. J. & Allen, T. M. Rate of biodistribution of STEALTH liposomes to tumor and skin: influence of liposome diameter and implications for toxicity and therapeutic activity. *Biochimica et biophysica acta* **1609**, 102–108 (2003).
- Dolgin, E. Animal testing alternatives come alive in US. *Nature medicine* **16**, 1348, <https://doi.org/10.1038/nm1210-1348> (2010).
- Pires, D. E., Blundell, T. L. & Ascher, D. B. pkCSM: Predicting Small-Molecule Pharmacokinetic and Toxicity Properties Using Graph-Based Signatures. *Journal of medicinal chemistry* **58**, 4066–4072, <https://doi.org/10.1021/acs.jmedchem.5b00104> (2015).
- Riviere, J. E., Scoglio, C., Sahneh, F. D. & Monteiro-Riviere, N. A. Computational approaches and metrics required for formulating biologically realistic nanomaterial pharmacokinetic models. *Computational Science and Discovery* **6**, 014005 (2013).
- Xia, X. R., Monteiro-Riviere, N. A. & Riviere, J. E. An index for characterization of nanomaterials in biological systems. *Nature nanotechnology* **5**, 671–675, <https://doi.org/10.1038/nnano.2010.164> (2010).
- Wolfgram, J. *et al.* The nano-plasma interface: Implications of the protein corona. *Colloids and surfaces. B, Biointerfaces* **124**, 17–24, <https://doi.org/10.1016/j.colsurfb.2014.02.035> (2014).
- Bachler, G., von Goetz, N. & Hungerbuhler, K. A physiologically based pharmacokinetic model for ionic silver and silver nanoparticles. *International journal of nanomedicine* **8**, 3365–3382, <https://doi.org/10.2147/IJN.S46624> (2013).
- Bachler, G., von Goetz, N. & Hungerbuhler, K. Using physiologically based pharmacokinetic (PBPK) modeling for dietary risk assessment of titanium dioxide (TiO₂) nanoparticles. *Nanotoxicology* **9**, 373–380, <https://doi.org/10.3109/17435390.2014.940404> (2015).
- Lankveld, D. P. *et al.* The kinetics of the tissue distribution of silver nanoparticles of different sizes. *Biomaterials* **31**, 8350–8361, <https://doi.org/10.1016/j.biomaterials.2010.07.045> (2010).
- Lee, H. A., Leavens, T. L., Mason, S. E., Monteiro-Riviere, N. A. & Riviere, J. E. Comparison of quantum dot biodistribution with a blood-flow-limited physiologically based pharmacokinetic model. *Nano letters* **9**, 794–799, <https://doi.org/10.1021/nl803481q> (2009).
- Lin, Z., Monteiro-Riviere, N. A. & Riviere, J. E. A physiologically based pharmacokinetic model for polyethylene glycol-coated gold nanoparticles of different sizes in adult mice. *Nanotoxicology* **10**, 162–172, <https://doi.org/10.3109/17435390.2015.1027314> (2016).
- Sweeney, L. M., MacCalman, L., Haber, L. T., Kuempel, E. D. & Tran, C. L. Bayesian evaluation of a physiologically-based pharmacokinetic (PBPK) model of long-term kinetics of metal nanoparticles in rats. *Regulatory toxicology and pharmacology: RTP* **73**, 151–163, <https://doi.org/10.1016/j.yrtph.2015.06.019> (2015).
- Gilkey, M. J. *et al.* Physiologically Based Pharmacokinetic Modeling of Fluorescently Labeled Block Copolymer Nanoparticles for Controlled Drug Delivery in Leukemia Therapy. *CPT: pharmacometrics & systems pharmacology* **4**, e00013, <https://doi.org/10.1002/psp4.13> (2015).
- Li, D. *et al.* Physiologically based pharmacokinetic modeling of polyethylene glycol-coated polyacrylamide nanoparticles in rats. *Nanotoxicology* **8**(Suppl 1), 128–137, <https://doi.org/10.3109/17435390.2013.863406> (2014).
- Li, M., Panagi, Z., Avgoustakis, K. & Reineke, J. Physiologically based pharmacokinetic modeling of PLGA nanoparticles with varied mPEG content. *International journal of nanomedicine* **7**, 1345–1356, <https://doi.org/10.2147/IJN.S23758> (2012).
- Tomalia, D. A. *et al.* A New Class of Polymers - Starburst-Dendritic Macromolecules. *Polym J* **17**, 117–132, <https://doi.org/10.1295/polymj.17.117> (1985).
- Kaminskas, L. M. *et al.* Doxorubicin-conjugated PEGylated dendrimers show similar tumoricidal activity but lower systemic toxicity when compared to PEGylated liposome and solution formulations in mouse and rat tumor models. *Molecular pharmaceutics* **9**, 422–432, <https://doi.org/10.1021/mp200522d> (2012).
- Ryan, G. M. *et al.* PEGylated polylysine dendrimers increase lymphatic exposure to doxorubicin when compared to PEGylated liposomal and solution formulations of doxorubicin. *Journal of controlled release: official journal of the Controlled Release Society* **172**, 128–136, <https://doi.org/10.1016/j.jconrel.2013.08.004> (2013).
- Wang, J. *et al.* The Role of Micelle Size in Tumor Accumulation, Penetration, and Treatment. *ACS nano* **9**, 7195–7206, <https://doi.org/10.1021/acs.nano.5b02017> (2015).
- Kaminskas, L. M. *et al.* Impact of surface derivatization of poly-L-lysine dendrimers with anionic arylsulfonate or succinate groups on intravenous pharmacokinetics and disposition. *Molecular pharmaceutics* **4**, 949–961, <https://doi.org/10.1021/mp070047s> (2007).
- Kaminskas, L. M. *et al.* The impact of molecular weight and PEG chain length on the systemic pharmacokinetics of PEGylated poly-L-lysine dendrimers. *Molecular pharmaceutics* **5**, 449–463, <https://doi.org/10.1021/mp7001208> (2008).
- Kaminskas, L. M. *et al.* Methotrexate-Conjugated PEGylated Dendrimers Show Differential Patterns of Deposition and Activity in Tumor-Burdened Lymph Nodes after Intravenous and Subcutaneous Administration in Rats. *Molecular pharmaceutics* **12**, 432–443, <https://doi.org/10.1021/Mp500531e> (2015).
- Pires, D. E. *et al.* Cutoff Scanning Matrix (CSM): structural classification and function prediction by protein inter-residue distance patterns. *BMC Genomics* **12**(Suppl 4), S12, <https://doi.org/10.1186/1471-2164-12-S4-S12> (2011).
- Pires, D. E. V., Kaminskas, L. M. & Ascher, D. B. Prediction and Optimization of Pharmacokinetic and Toxicity Properties of the Ligand. *Methods Mol Biol* **1762**, 271–284, https://doi.org/10.1007/978-1-4939-7756-7_14 (2018).
- Pires, D. E. & Ascher, D. B. CSM-lig: a web server for assessing and comparing protein-small molecule affinities. *Nucleic Acids Res* **44**, W557–561, <https://doi.org/10.1093/nar/gkw390> (2016).
- Pires, D. E. & Ascher, D. B. mCSM-AB: a web server for predicting antibody-antigen affinity changes upon mutation with graph-based signatures. *Nucleic Acids Res* **44**, W469–473, <https://doi.org/10.1093/nar/gkw458> (2016).
- Pires, D. E. V. & Ascher, D. B. mCSM-NA: predicting the effects of mutations on protein-nucleic acids interactions. *Nucleic Acids Res* **45**, W241–W246, <https://doi.org/10.1093/nar/gkx236> (2017).
- Pires, D. E., Ascher, D. B. & Blundell, T. L. DUET: a server for predicting effects of mutations on protein stability using an integrated computational approach. *Nucleic Acids Res* **42**, W314–319, <https://doi.org/10.1093/nar/gku411> (2014).
- Pires, D. E., Ascher, D. B. & Blundell, T. L. mCSM: predicting the effects of mutations in proteins using graph-based signatures. *Bioinformatics* **30**, 335–342, <https://doi.org/10.1093/bioinformatics/btt691> (2014).
- Pires, D. E., Blundell, T. L. & Ascher, D. B. mCSM-lig: quantifying the effects of mutations on protein-small molecule affinity in genetic disease and emergence of drug resistance. *Sci Rep* **6**, 29575, <https://doi.org/10.1038/srep29575> (2016).
- Pires, D. E., Chen, J., Blundell, T. L. & Ascher, D. B. *In silico* functional dissection of saturation mutagenesis: Interpreting the relationship between phenotypes and changes in protein stability, interactions and activity. *Sci Rep* **6**, 19848, <https://doi.org/10.1038/srep19848> (2016).
- Rodrigues, C. H., Pires, D. E. & Ascher, D. B. DynaMut: predicting the impact of mutations on protein conformation, flexibility and stability. *Nucleic Acids Res* **46**, W350–W355, <https://doi.org/10.1093/nar/gky300> (2018).
- Rodrigues, C. H. M., Myung, Y., Pires, D. E. V. & Ascher, D. B. mCSM-PP2: predicting the effects of mutations on protein-protein interactions. *Nucleic Acids Res* **47**, W338–W344, <https://doi.org/10.1093/nar/gkz383> (2019).

37. Rodrigues, C. H., Ascher, D. B. & Pires, D. E. Kinact: a computational approach for predicting activating missense mutations in protein kinases. *Nucleic Acids Res* **46**, W127–W132, <https://doi.org/10.1093/nar/gky375> (2018).
38. Hall, M. *et al.* The WEKA data mining software: an update. *SIGKDD Explor. Newsl.* **11**, 10–18, <https://doi.org/10.1145/1656274.1656278> (2009).
39. Breiman, L. Random Forests. *Machine Learning* **45**, 5–32, <https://doi.org/10.1023/A:1010933404324> (2001).

Acknowledgements

L.M.K. was funded by an NHMRC Career Development Fellowship. D.B.A. and D.E.V.P. were funded by a Newton Fund RCUK-CONFAP Grant awarded by The Medical Research Council (MRC) and Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) (MR/M026302/1, APQ-00828-15). D.E.V.P. received support from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) (409780/2016-2), Brazil. DBA was supported by a C.J. Martin Research Fellowship from the National Health and Medical Research Council of Australia (APP1072476) and the Jack Brockhoff Foundation (JBF 4186, 2016). This work was supported in part by the Victorian Government's OIS Program.

Author contributions

D.B.A. designed and directed the project. L.M.K. performed the data curation. D.E.V.P. and D.B.A. performed the analysis and model development. All authors interpreted the results and wrote the manuscript. All authors have given approval to the final version of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-019-51789-3>.

Correspondence and requests for materials should be addressed to L.M.K. or D.B.A.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019