ISSN 0976 – 9080

BIOMIRROR - An Open Access Journal

www.bmjournal.in BM/Vol.1/September 2010/bm-0808131710

Computational approaches for Oral bioavailability prediction: An overview

Rajnish Kumar*, Anju Sharma, Pritish Kumar Varadwaj

Indian Institute of Information Technology, Allahabad, Uttar Pradesh, India

Abstract

High oral bioavailability is among the most important consideration during drug development process. Oral-bioavailability is usually determined in the pre-clinical stage of drug development process. It was found that 30% of drugs fail during the drug discovery process. Therefore there is a need of a robust and accurate computational model which can predict the oral bioavailability of compounds without carrying out any experiments. There exists a plethora of studies to predict oral bioavailability which indicates that it is incredibly rich area of research. Various attempts in estimating oral bioavailability are reported in literature belonging to different categories. In this article computational oral bioavailability prediction approaches are discussed.

Keywords: Oral bioavailability, Prediction, Drugs, Computational methods

Statistical approaches to predict oral bioavailability

Lipsinki's 'Rule-of-Five' [1] pioneered the in-silico prediction of oral bioavailability. Since Lipinski et al's influential findings, several groups have studied this issue and tried to decipher other factors that may also be important for oral bioavailability. Navia [2] postulated that to pass across the permeable membrane drug should have desirable molecular flexibility. In another work Hirschmann [3] found that property of water to complex amide bond has negative effect on oral bioavailability. Veber et al [4] studied rat oral bioavailability data, acquired by Glaxo SmithKline, for over 1100 drug candidates. They found that reduced molecular flexibility and low Polar Surface Area (PSA) are important predictors of good oral bioavailability.

*** Corresponding author: Mr.Rajnish Kumar ***
Indian Institute of Information Technology,Allahabad, Uttar Pradesh, India.
Email: 12.rajnish@gmail.com

Received September 09, 2010; **Accepted** September 13, 2010; **Published** September 17, 2010

Citation: Rajnish Kumar*, Anju Sharma, Pritish Kumar Varadwaj (2010-Computational approaches for Oral bioavailability prediction: An overview) BIOMIRROR : 1-7/ bm-0808131710

They suggested that the molecule having less than 10 rotatable bonds and its PSA is less than 140 Å^2 will have good oral bioavailability, especially in rats exceeding 20%. In a recent study Martin et al., [5] tried to map bulkiness and various polarity factors

of lipophillic drugs to OB and concluded that bioavailability depends on molecular weight. It is maximal at 31-35 $(J/cm^3)^{1/2}$. Lu et al. [6] studied the relationship between PSA and rotatable bond count to predict oral bioavailability using 434 compounds and compared their results with Veber et al's. The results were in support of Veber's finding. They also found that the correlations were dependent on calculation methods. Therefore Lu et al, suggested care should be taken before using any generalization. According to Manuel et al [7], factors like partition coefficient, conformational flexibility, molecular weight etc can act as guiding principles to predict oral bioavailability but these are not only ones. In a novel attempt to estimate the bioavailability Martin [8] recently formulated "a bioavailability score" (ABS) on the basis of molecular properties like molecular charge, Lipsinki's Rule of Five, PSA. They used a diverse set of 553 compounds with their known bioavailability rat data. ABS was also examined for human data of 449 compounds. They found that Rule of Five is able to identify compounds with poor availability that are either neutral or positively charged but hardly predict anionic compounds. On the other hand, PSA rule can predict anionic compounds, not for neutral and cationic ones. Vieth.et al [9] analyzed 1729 marketed drugs and found that drugs given orally are light weighted and contains very less rotatable bonds and hydrogen bond donor/acceptor than drugs which are given by other routes. Again, this was in good coincidence with Veber et al's findings. Sietsema [10] studied oral bioavailability of hundreds of drugs. They also carried out interspecies comparison and observed that there was no correlation of absolute oral bioavailability between species though it can be used to infer relative comparisons. Hou et al [11] studied the correlation between several different molecular properties and human oral bioavailability data using dataset of 768 chemical compounds. He postulated that a simple rule which is based on molecular properties cannot be used to predict oral bioavailability with good confidence.

Mechanism-Based Oral Bioavailability Models

These models provide mechanistic details and associated underlying hypotheses. Somogyi et al [12] tried to predict the bioavailability of drugs having high first-pass effects by generating a plot between oral clearance and reciprocal of bioavailability. Since a drug with high hepatic clearance has low bioavailability therefore they observed a straight line with intercept of unity having a slope of reciprocal of hepatic blood flow. Usansky et al. [13] developed a model for absorption-disposition which predicts human oral bioavailability using data from in-vitro permeability and from computationally predicted permeability as well. For 49 out of 51 compounds in study, the residuals between predicted and experimental bioavailability values ranged from 17% to 22%. Mahmood [14] calculated oral bioavailability by plotting oral bioavailability against renal clearance after a single oral dose. The study was based on the analysis of only 8 drugs therefore is not

reliable for the prediction of absolute bioavailability. Obata et al [15] developed a theoretical passive absorption models (TPAM) using a data set of 258 compounds with observed oral bioavailability values. logD at pH 6.0, intrinsic logP, pKa and molecular weight were calculated from the chemical structure. The TPAM predicted the bioavailability values with RMSE of 155 to 21% and a correlation coefficient of 0.78 to 0.88. The possibility of over-learning was low because there were only four coefficients in models that were optimized by fitting with hundreds of bioavailability values. A physiologically based model to predict human bioavailability was developed by William et al [16] which is based on the permeability-limited absorption. Model parameters were optimized using 126 compounds. The RMSD was observed to be 7% for passively absorbed compounds. Arun [17] developed a graphical approach based on combined measurements of CaCo-2 flux and microsomal stability to predict bioavailability by examining 21 drugs. These drug candidates had a wide range of oral bioavailability values. He used reference plots to make a prediction i.e. classifies a compound as 0-20%, 20-50%, or 50- 100% bioavailable.

QSAR/QSPR Approaches to Model Bioavailability

Various efforts have been made in recent years to model oral bioavailability using descriptor based QSAR/QSPR approaches. Both linear and nonlinear learning methods have been applied to develop quantitative model for the prediction of bioavailability. Yoshida et al [18] trained a QSAR model based on 3 descriptors related to distribution coefficient (logD at pH 7.4-6.5) as inputs for physicochemical properties and 15 structural descriptors indicating the presence/ absence of certain functional groups that are most likely to be involved in metabolic reactions. This approach used "fuzzy adaptive least squares" and can classify drugs with an accuracy of 60%. Important finding was that acids generally had better bioavailability than bases, with neutral compounds in between. Another effort based on classification using SIMCA approach achieved similar success [19]. Hologram QSAR model was developed by Tiago et al [20] for a set of 250 structurally diverse molecules using atoms, bond, connection and chirality as fragment distinction. HQSAR was tested on test set of 52 molecules. Prediction of oral bioavailability was found to be in good agreement with experimental values (q^2 = 0.70 and r^2 = 0.93). Andrew et al [21] analyzed 591 compounds with human oral bioavailability using substructure count descriptors, regression and recursive partitioning and gave a QSBR model. Model achieved prediction with RMSD of ~18%. As compared to Lipsinki's Rule of five false-negative predictions were reduced from 5% to 3% while false positive prediction decreased from 78% to 53%. Andrew regression model suffers from problem of overfitting. Despite large degree of experimental error QSBR model was reasonably used. Bai et al [22] used a classification regression trees to analyze human oral bioavailability and absorption

data for a set of 1261 structures (899 as training set and 362 as test set). Compounds were divided into six classes. On two test sets, model achieved correct classification rates ranging from 79% to 86%. Zmuidinavivius et al [23] analyzed 1000 drugs like compounds with known human intestinal absorption values using recursive partitioning and were able to achieve 15% false positive and 3% false negative classification rates. Joseph et al [24] developed QSPR for 169 compounds using eight descriptors representing hydrophobic, steric, electronic and constituent parameters of drugs. A correlation of 0.72 was achieved for test set bioavailability prediction when compared with literature values. Klopman et al [25] developed a model based on a modified group contribution method using the computer automated structure evaluation (CASE) program for the dataset of 417 compounds. The model was able to predict the percentage of drug absorbed with an R^2 of 0.79 and a standard deviation of 12.3% for the compounds from training set. The standard deviation for an extended test set of 50 drugs was 12.3%.

Genetic Programming based prediction of oral bioavailability

Genetic programming (GP) uses all features in the dataset and automatically performs strong and efficient feature selection. Francesco et al [26] carried out experimental study of quantitative prediction of bioavailability of 241 molecules. Four different versions of GP, which differs in the fitness function and in set of terminal symbols used, have been tested. The version which optimizes both the root means square error and correlation coefficient on training set has shown the best generalization capability. On similar lines Sara et al [27] proposed operator equalization as a new bloat control technique for genetic programming. This technique proved to be more efficient and showed more predictability than previous methods. Bains et al [28] also used genetic programming for predicting bioavailability. The results showed a slight improvement as compared with Yoshida [18] approach, although a direct comparison is difficult owing to a different selection of bioavailability range of the classes. Another application of genetic algorithm to build QSPR model for prediction of human absolute oral bioavailability was given by Junmei et al [29], score improved the correlation coefficient and reduced the standard error significantly. The key fragments that increase or reduce pharmacokinetic properties were also identified. Pintore [30] used adaptive fuzzy partitioning (AFP) to predict oral bioavailability. The best molecular descriptors were selected using genetic algorithm and stepwise methods further they were classified using Self organizing maps (SOMs). In 75% of cases molecules were correctly classified in the right bioavailability of class.

Artificial Neural Network (ANN) is widely used for ADMET parameters estimation and there are many studies in which a Bayesian based Artificial Neural Networks (BRANN) have been applied. Turner et al [31] analyzed human bioavailability data for 167 compounds using Artificial Neural Network

(ANN). Model was trained with 137 compounds and tested with test set of 15 compounds followed by further validation of another 15 compounds. The model was able to classify between low and high bioavailability. Fröhlich et al [32] recently studied the use of Kernel methods (SVM) for assessing the problem of bioavailability predictions. Their approach was based on the estimation of similarity between different molecules showing similar biological behavior. Lu et al [33] analyzed 169 compounds (113 in training set and 56 in test set) using Support Vector Machine (SVM) and the regression methods. These were further implemented in comprehensive descriptors for structural and statistical analysis (CODESSA). Both linear and non-linear models gave satisfactory prediction results.

References

1. Lipsinki CA, Lomabardo F, Dominy BW, Feeney PJ, Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv. Drug Delivery Rev. 23: 4-25 (1997)

2. Navia MA, Chaturvedi PR, Design Principles for Orally Bioavailable drugs. Drug Discovery Today, 1:179-189 (1996)

3. Hirschmann R, Peptide research a means to further biological and chemical understanding. In Peptides 1996. Proceedings of the 24th European peptide Symposium.

4. Veber DF , Johnson SR , Cheng HY , Smith BR , Ward KW , Kopple KD . Molecular properties that influence the oral bioavailability of drug candidates. J Med Chem, 45: 2615 - 2623 (2002)

5. Martin TK and Yvonne A, Influence of molecular properties on oral bioavailability of lipophilic drugs-Mapping of bulkiness and different measures of polarity, Pharm. Dev. And Tech., 14(3): 312-320 (2009)

6. Lu JJ, Crimin K and Goodwin JT,. Influence of molecular flexibility and polar surface area metrics on oral bioavailability in the rat. J Med Chem, 47 : 6104 – 6107 (2004)

7. Manuel AN and Pravin RC, Design principles for orally bioavailable drugs, Drug Discovery Today, 1(5): 179-189 (1999)

8. Martin YC, A bioavailability score. J Med Chem, 48 : 3164 - 3170 (2005)

9. Vieth M, Siegel MG and Higgs RE, Characteristic physical properties and structural fragments of marketed oral drugs. J Med Chem, 47 : 224 - 232 (2004)

10. Sietsema WK, The absolute oral bioavailability of selected drugs. J.Pharmacol Toxicol Methods, 44(1): 235-49 (2000)

11. Hou T, Wang J, Zhang W and Xu X, ADME evaluation in Drug Discovery.6. Can Oral bioavailability in humans be predicted

by simple molecular property-based rules, J. Chem. Inf. Model, 47: 460-463 (2007)

12. Somogyi A, Eichelbaum M and Gugler R, Prediction of bioavailability for drugs with a high first-pass effect using oral clearance data, Eur. J. of Clinical Pharm. 22(1): 85-90 (2006)

13. Usansky HH and Sinko PJ, Estimating human drug oral absorption kinetics from Caco-2 permeability using an absorption-disposition model: model development and evaluation and derivation of analytical solutions for k(a) and F(a). J Pharmacol Exp Ther., 314 : 391 – 399 (2005)

14. Mahmood I, Prediction of absolute bioavailability for drugs using oral and renal clearance following a single dose: A critical view, Biopharm. and Drug Dispos. 18(6): 465-473 (1997)

15. Obata K, Sugano K and Saitoh R, Prediction of oral drug absorption in humans by theoretical passive absorption model. Int J Pharm. 293: 183 – 192 (2005)

16. Willmann S, Schmitt W, Keldenich J, Lippert J and Dressman JB, A physiological model for the estimation of the fraction dose absorbed in humans. J Med Chem. 47: 4022 - 4031 (2004)

17. Arun KM, Thomas NT and Hwang KK, Graphical model for estimating oral bioavailability of drugs in humans and other species from there Caco-2 permeability and in vitro liver enzyme metabolic stability rates, J. Med. Chem. 45(2):304-311 (2002)

18. Yoshida F. and Topliss, JG, QSAR model for drug human oral bioavailability, J. Med. Chem. 43: 2575–2585 (2000)

19. Podlogar BL, Muegge I. and Brice LJ, Computational methods to estimate drug development parameters. Curr. Opin. Drug Disc. Dev. 4: 102–109 (2001)

20. Tiago LM, Carlos AM and Adriano DA, Hologram QSAR model for the prediction of human oral bioavailability, Bioorganic and Medicinal Chemistry, 15(24): 7738-7745 (2007)

21. Andrews CW, Bennett L. and Yu LX, Predicting human oral bioavailability of a compound: development of a novel quantitative structure-bioavailability relationship. Pharm. Res. 17: 639–644 (2000)

22. Bai JP, Utis A and Crippen G. Use of classification regression tree in predicting oral absorption in humans. J Chem Inf Comput Sci., 44 : 2061 - 2069 (2004)

23. Zmuidinavicius D, Didziapetris R, Japertas P, Avdeef A and Petrauskas A, Classification structure-activity relations (C-SAR) in prediction of human intestinal absorption. J Pharm Sci., 92: 621 – 633 (2003)

24. Joseph VT, Beverly DG and Snezana AK, Prediction of drug bioavailability based on molecular structure, Analytica Chimica Acta, 485(1): 89-102 (2003)

25. Klopman G, Stefan LR and Saiakhov RD, ADME evaluation, II: a computer model for the prediction of intestinal absorption in humans. Eur J Pharm Sci. 17: 253 – 263 (2002)

26. Francesco A, Stefano L, Enza M and Leonardo V, Genetic programming for Human bioavailability of drugs, GECCO'06, 255-262 (2006)

27. Sara S and Leonardo V, Operator equalization, bloat and overfitting: a study on humab bioavailability prediction, ACM, 1115-1122 (2009)

28. Bains W, Gilbert R, Sviridenko L, Gascon JM, Scoffin R, Birchall K, Harvey I. and Caldwell J, Evolutionary computational methods to predict oral bioavailability QSPRs. Curr. Opin. Drug Disc. Dev. **5**: 44–51 (2002)

29. Junmei W, George K, Xie XQ, Chengde W and George H, Genetic Algorithm-Optimized QSPR Models for Bioavailability, Protein Binding, and Urinary Excretion, J. Chem. Inf. Model., 46 (6): 2674–2683 (2006)

30. Pintore M, Van de Waterbeemd H, Piclin N. and Chrétien JR, Prediction of Oral Bioavailability by Adaptive Fuzzy Partitioning, Eur. J. Med. Chem. (in the press).

31. Turner JV, Maddalena DJ and Agatonovic-Kustrin S, Bioavailability prediction based on molecular structure for a diverse series of drugs. Pharm Res., 21 : 68 - 82 (2004)

32. Fröhlich JW, Sieker F and Zell A, Kernel Functions for Attributed Molecular Graphs - A New Similarity Based Approach To ADME Prediction in Classification and Regression, QSAR & Combinatorial Science, 2005.

33. Liu HX , Hu RJ , Zhang RS , et al . The prediction of human oral absorption for diffusion rate-limited drugs based on heuristic method and support vector machine. J Compute Aided Mol Des, 19: 33-46 (2005)