# Rational Computer-Aided Design of Biomaterials

W J Welsh[1] , V Kholodovych[1], D Knight[2], J Kohn[2]
[1]University of Medicine & Dentistry of NJ, Piscataway, NJ, US
[2]Rutgers, The State University of NJ, New Brunswick, NJ, US

## Introduction

Identifying specific correlations between cellular response and materials structure and properties is time-consuming and expensive when a large number of biomaterials require evaluation. Here we present a surrogate (semi-empirical) computational model for prediction of cellular response to the surfaces of biodegradable polymers that have been designed for tissue engineering applications. The predictions of our model, when compared with experimental results, showed excellent agreement. The model was determined by fitting experimental data for a series of 62 polyarylates to a small number of polymer structure-based "molecular descriptors" using the technique of Partial Least Squares (PLS) regression. Quantitative predictions of cellular response to six polymers (untested prior to model building) concurred with experiment within 15.8% on average.

## Materials and Methods

Combinatorial chemistry techniques were adapted to prepare a series of structurally related polyarylates derived from monomers consisting of a tyrosine-derived diphenol and a diacid [1]. All possible combinations with the available 14 diphenols and 8 diacids yield 112 structurally distinct, but closely related, polymers. Of the possible 112 polyarylates, 62 compositions were selected randomly for testing prior to modeling. Following model building, six compositions that did not belong to the 62 polymer training set were tested for cellular response in separate experiments (same protocol). These additional six polymers were chosen to probe the accuracy of model predictions over their full range. Two of them were predicted to yield high values of FRLF NMA, two were predicted to yield middling values and two very low values.

Metabolic activity of cells as a function of the specific polymer substratum was evaluated for Fetal Rat Lung Fibroblast (RFL-6). The experimental data were converted to values of Normalized Metabolic Activity (NMA) [1].

Model compounds were constructed for each polymer using the Molecular Operating Environment (MOE) molecular modelling software (www.chemcomp.com). Fifteen empirical descriptors for each polymer were generated using the Dragon Software (www.disat.unimib.it/chm) and 104 molecular topological descriptors were calculated using MOE. Partial Least Squares (PLS) analysis is a commonly employed regression technique in drug discovery for building linear QSPR models. PLS is noted for its speed, general reliability, and ease of interpretation [2].

## Results and Discussion

PLS models were developed based on the experimental NMA cellular response data. The final PLS model was statistically self-consistent ($r^2$=0.62) and internally predictive ($r_{CV}^2$=0.56). Corresponding values of the PLS-predicted versus experimental NMA are plotted in Fig 1.

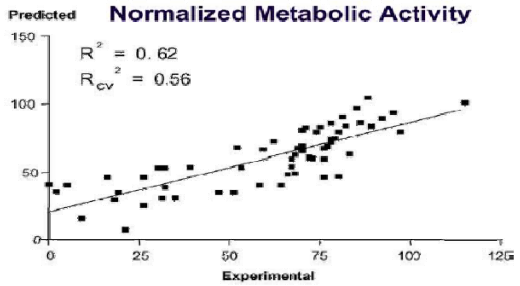


Fig.1 PLS model: Predicted vs Experimental NMA values

The PLS model was deployed to predict cellular response in the presence of each of the remaining 50 virtual polymers. Six of these polymers were chosen for experimental verification of these predictions. Each prediction fell within the experimental error except for one case (HTH methyladipate). It is worth noting that the average percent error of prediction (15.8%) is much less than the experimental average percent standard deviation (23.1%) and is nearly an order of magnitude less than the total variation in the set of 6 test polymers (145%). The model discriminated between the highest and lowest performers to within the level of experimental error and, indeed, has done so prior to any experimentation [3, 4].

| Name | Predicted NMA (% TCPS) | Measured NMA (% TCPS) |
|---|---|---|
| HTH Methyladipate | 33.67 | 63.7± 12.1 |
| DTiB Adipate | 40.91 | 41.4 ± 7.9 |
| DTiB Diglycolate | 54.96 | 62.6 ± 11.9 |
| HTH Glutarate | 59.45 | 53.2 ± 10.1 |
| HTE Adipate | 69.68 | 67.1 ± 12.7 |
| HTE Diglycolate | 82.62 | 101.5 ± 19.3 |

## Summary/Conclusions

The present study demonstrates that surrogate QSPR-based models can be employed to offer guidance and direction in the design, selection, and optimization of novel biorelevant materials. Since computed polymer descriptors are less expensive to obtain than *in vitro* or *in vivo* measurements, the use of a computational modeling approach can significantly reduce the costs and labor associated with identifying high-performance biomaterials for specific applications.

## References

1. Brocchini S. et al. Biomed. Mater. Res., 1998, 42, 66.
2. Wold S. Methods and principles in medicinal chemistry. Weinheim, Germany: Verlag; 1995.
3 Kholodovych V. et al. Polymer, 2004, 45, 7367-7379
4. Smith J. et al.Polymer (in press)

## Acknowledgment

This work was supported in part by "RESBIO - The National Resource for Polymeric Biomaterials" funded under NIH grant P41 EB001046-01A1.