Discovery of Polymers for Plasmid DNA Delivery using Combinatorial Synthesis and Cheminformatics Modeling

Kaushal Rege¹, Thrimoorthy Potta¹, Zhuo Zhen², Taraka Sai Pavan Grandhi³, Matthew Christensen¹, James Ramos³, Curt M. Breneman²

¹Chemical Engineering, Arizona State University, Tempe, AZ 85287-6106; ²Department of Chemistry and Chemical Biology, Rensselaer Polytechnic Institute, Troy, NY 12180; ³Harrington Biomedical Engineering, Arizona State University, Tempe, AZ 85287-6106

kaushal.rege@asu.edu

ABSTRACT SUMMARY
A library of fifty-six novel aminoglycoside antibiotic-based polymers was developed using combinational synthesis methods. Parallel screening of this library led to the identification of several polymers which exhibited higher transgene expression and lower toxicities compared to the current polymeric standard pEI in prostate and pancreatic cancer cells. Quantitative Structure Activity Relationship (QSAR)-based cheminformatics modeling was employed to correlate transgene expression efficacies with physicochemical properties of polymers. QSAR models were able to predict the transgene expression efficacy of polymers not included in model training. Model interpretation using molecular descriptors indicated that aminoglycoside monomer size, cross-linker length, location of the hydrophilic oxygen atom, and presence of aromatic rings significantly affect the polymer transgene expression. Our results demonstrate the power of using combinatorial and cheminformatics methods for rapid discovery and mechanistic elucidation of polymers for delivery of plasmid DNA for transgene expression

INTRODUCTION
Polymeric systems have been widely investigated for plasmid DNA delivery¹ due to safety concerns with viral vectors. However, low efficacies and high cytotoxicities associated with several polymeric vectors (e.g. pEI) necessitate new predictive approaches that facilitate rapid discovery of effective polymeric vehicles for transgene expression. In this study, we generated a novel aminoglycoside-based polymer library using combinational synthesis methods, and rapidly screened the library for transgene expression efficacy. We used this chemical information for generating predictive QSAE models to correlate physicochemical properties of polymers to their efficacies.

EXPERIMENTAL METHODS
Ring opening polymerization of amines from aminoglycosides and epoxides from diglycidyl ethers was used to generate the polymer library. Purified polymers were characterized using ¹H-NMR and FT-IR spectroscopy. Physicochemical properties such as polymer molecular weights, polyplex size and zeta potential were measured. Transgene expression efficacies in PC3, 22Rv1 prostate cancer cells and MiaPaCa-2 pancreatic cancer cells were determined using luciferase and GFP expression. QSAR models were generated using physicochemical molecular descriptors, K-fold cross-Validation, Model Evaluation and Y-Scramble Technique to elucidate the role of polymer physicochemical properties on transgene expression efficacies.

RESULTS AND DISCUSSION
¹H-NMR and FT-IR studies confirmed polymer formation. Polymers effectively condensed plasmid DNA to stable polyplexes of size 90-150nm. Parallel screening of the aminoglycoside polymer library using PC3 (Figure 1) and MiaPaCa-2 resulted in the identification of several lead polymers which exhibited higher luciferase expression than 25 kDa poly(ethylene imine) or pEI. Seven polymers that exhibited highest levels of transgene expression were used in subsequent investigations. Cytotoxicity studies polymers revealed that these polymers show negligible toxicities under the doses tested (0.4 – 40 µg/mL).
Flow cytometry / dose response studies indicated that up to 23% cells were positive for GFP in case of lead polymers, which was significantly higher than what was observed for pEI.

Quantitative Structure Activity Relationship Models were generated using Support Vector Regression (SVR) for a set of 33 polymers from the library. Descriptor (feature) selection, in concert with cross validation methods, was used to as used to prevent overfitting of the model. The SVR-based QSAR model had a squared Pearson’s correlation coefficient close to 1.0 and a coefficient of determination of 0.86 for our external test set. Predictions for a test set of polymers not included in the training of the model, showed excellent agreement with the experimental values, indicating the predictive power of this approach. Y-scrambling analyses demonstrated that the regression was not over-trained and was robust, and therefore elegant. Interpretation of the descriptors in the predictive QSAR model indicated that larger size of aminoglycoside monomers, shorter chain lengths of diglycidyl ether crosslinkers, and presence of larger numbers oxygen atoms and ring structures, enhanced transgene delivery.

CONCLUSION

A library of fifty-six polymers was synthesized and screened leading to the identification of several lead polymers which exhibited higher transgene expression compared to the commercially available polymeric standard pEI. Support Vector Regression based QSAR modeling was carried out with a set of polymers from the model, and the predictive ability of the QSAR model showed excellent agreement with values for a test set not included in the model training. QSAR model interpretation indicated that size, presence of hydrophilic oxygen atoms and shorter hydrocarbon chain between ether oxygen atoms in the cross-linkers are favorable for efficient transgene delivery. To our knowledge this is the first report that describes (1) combinatorial synthesis of an aminoglycoside based library for transgene delivery, and (2) uses sophisticated cheminformatics modeling to elucidate physicochemical properties that enhance polymer-mediated transgene expression. This powerful approach can have a transformative impact on the rational design of non-viral materials for nucleic acid delivery.

REFERENCES


ACKNOWLEDGEMENTS

We are grateful to the NIH/NIGMS (Grant 1R01GM093229-01A1) and the Arizona Biomedical Research Commission (ABRC; Grant 1009) for financial support of this study.