Theoretical Calculation of Nonpolar Surface Areas of Glycam with Implicit Solvent Methods and Its Application in Glycomics Analysis.

Bhumika Jayee\textsuperscript{1}, Rabin Dhakal\textsuperscript{2}, Reed Nieman\textsuperscript{3}, Daniel Valente\textsuperscript{4}, Thiago M. Cardozo\textsuperscript{4}, Amna Aqdas\textsuperscript{1}, Wenjing Peng\textsuperscript{1}, Adelia Aquino\textsuperscript{2}, Hans Lischka\textsuperscript{1}, Yehia Mechref\textsuperscript{1}, Hanna Moussa\textsuperscript{2}

\textsuperscript{1}Department of Chemistry and Biochemistry, Texas Tech University, Lubbock, TX, USA
\textsuperscript{2}Department of Mechanical Engineering, Texas Tech University, Lubbock, TX, USA
\textsuperscript{3}Department of Chemistry and Chemical Biology, University of New Mexico, Albuquerque, NM
\textsuperscript{4}Instituto de Química, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, Brazil

\textbf{Abstract}

Glycomics is an interesting research area to comprehensively understand the biological attributes of glycopolymers in many important biological functions. The most efficient approach to analyze glycans has been mass spectrometry (MS) due to its high sensitivity and ability to offer structural information. However, it remains analytically challenging due to the microheterogeneity of glycans and the limited ability to identify glycan isomers compounded by the lack of standards. In this context, a property of interest is the nonpolar surface area (NPSA) as it has been shown to correlate quite well to glycan retention times. Hence, accurate theoretical estimation of NPSA through a well-established and fast quantum chemical protocol could provide an invaluable tool in the analysis of biological samples. In this research, a procedure to obtain such estimates based on implicit solvent models is proposed and tested on glycan isomers. For the analysis, the quantum chemical calculations were performed using the program TURBOMOLE. NPSA values thus calculated for the optimized structural isomers were correlated to experimental retention times and demonstrate a coefficient of determination (R$^2$) of 0.989, suggesting a strong linear correlation. Both larger NPSA and larger retention times (in a C18 column) suggest larger nonpolar interactions, providing a possible physical basis for the observed correlation between these two variables. This research gives a novel method of theoretical calculation of NPSA through a simple and well-defined scheme using reliable and fast quantum chemical methods.

\textbf{Introduction}

Glycomics is the study of glycans that focuses on structures and function of carbohydrates and on glycoinformation distribution at cellular, tissues and organisms levels. The structure of glycans determine the carbohydrate binding protein partners to which a glycoprotein binds. Carbohydrate binding domains are common in cell surface and secreted proteins. In order to understand fully the structure-function relationships for glycoproteins, analysis of the glycan structure is unavoidable. For the analysis of glycans structure, various method has been used such as derivatization (e.g. permethylation) and separation techniques (e.g. liquid chromatography) preceding a mass spectrometry analysis. But there is challenge in analysis of glycans structure due to the microheterogeneity of its structures and these method have limited ability to identify glycan isomers. So, this research is focused on finding a novel techniques that is useful in structural analysis of glycans.

\textbf{Methods}

The initial structure for glycans were generated using GLYCAM-web Carbohydrate Builder tool (http://glycam.org/), where the glycotic glycan structure was created to a series of conformers with consistent structural characteristics. The PBE and SV basis set was used to speed-up the calculations. The Conductor-like Screening Model (COSMO), an implicit solvent model, was used as solvent model to simulate solvent effects as well as to calculate the NPSA. Each point on the COSMO solvent-accessible surface has associated charge and area values, enabling the calculation of the NPSA by simply summing the area of all points with charges between chosen threshold values. The TURBOMOLE electronic structure program package was utilized for all quantum chemistry calculations.

\textbf{Objectives}

1. To provide accurate theoretical estimation of NPSA through a well-established and fast quantum chemical protocol which could provide an invaluable tool in the analysis of biological samples.
2. To propose a procedure to obtain NPSA estimates based on implicit solvent models.
3. To test the proposed procedure to obtain NPSA estimates based on implicit solvent models on glycan isomers, providing needed structural information for the analysis of experimental results.

\textbf{Results}

\begin{figure}[ht]
\centering
\includegraphics[width=\textwidth]{figure1.png}
\caption{Dextran linear glycan, Glc12, optimized using the PBE-D3/SV method in the COSMO/acetonitrile environment. Representation a) shows the polarization charges, and b) the polar atoms in red, nonpolar atoms contributing to the NPSA in blue, and buried nonpolar atoms that do not contribute to the NPSA in yellow.}
\end{figure}

\begin{figure}[ht]
\centering
\includegraphics[width=\textwidth]{figure2.png}
\caption{Linear regression correlating experimentally determined retention times with NPSA estimated using the PBE-D3/SV approach. The results for the three investigated conformers are shown in a) gg, b) tg, and c) gt.}
\end{figure}

\begin{figure}[ht]
\centering
\includegraphics[width=\textwidth]{figure3.png}
\caption{Binominary standard glycan structures optimized using the PBE-D3/SV method in the COSMO/acetonitrile environment. Representations a) and c) show the polarization charges, and b) and d) the polar atoms in red, nonpolar atoms contributing to the NPSA in blue, and buried nonpolar atoms that do not contribute to the NPSA in yellow. The core-fucosylated structure is shown in a) and b), while the branch-fucosylated structure is shown in c) and d). Symbolic representations: blue square – N-Acetylglucosamine (GlcNAc), green circle – mannose, yellow circle – galactose, red triangle – fucose.}
\end{figure}

\begin{figure}[ht]
\centering
\includegraphics[width=\textwidth]{figure4.png}
\caption{Linear regression correlating experimentally determined retention times with NPSA calculated with the PBE-D3/SV approach for the core- and branch-fucosylated isomers and the 2,3- and 2,6-sulfated isomers using the PBE-D3/SV approach.}
\end{figure}

\textbf{Conclusion}

- The method proposed here is a general strategy to calculate the NPSA of glycans.
- It is based on induced surface charges (NPSA-ISC).
- It is calculated as the sum of the surface values of those surface elements with a surface charge $|q|$ less than a threshold.
- The result obtained also shows strong correlation of the NPSA values calculated, with measured retention times.

\textbf{References}