

*Application of In Silico Chromatography to Identify Bacteriohopanepolyol Diastereomers*  
Scott Hardy, Libby Allcock, Claire Vassallo, Alex Hamilton, Tim Nichol, Rachel Schwartz-Narbonne

### *Abstract*

Elucidating the role of marine anammox is key to investigations of the geological nitrogen cycle. A recently identified bacteriohopanepolyol (BHP) lipid biomarker, a diastereomer of bacteriohopanetetrol (bacteriohopane-17 $\beta$ ,21 $\beta$ (H), 22*R*, 32*R*, 33*R*, 34*S*, 35-tetrol, abbreviated BHT-34*S*) with unknown stereochemistry, holds promise as a diagnostic lipid marker for the marine anammox bacteria '*Ca. Scalindua*'. However, its use is limited by the unknown stereochemistry. Building on previous work by Allcock *et al.*, we further evaluated an analytical modelling approach based upon computational chemistry. We related gas chromatography (GC) retention time to polarizability calculated using density functional theory (DFT) calculations using known BHP structures and retention times. These improved computational methods provide a better fit than initial models. Other quantum descriptors such as frontier molecular orbital (FMO) energies, HOMO-LUMO gap and dipole moment were also explored. Additionally, the molecule-column binding energies were calculated, but with lower-cost computational methods; these did not provide good data fits.

### *Introduction*

Nitrogen is an essential element for life. Lack of bioavailable nitrogen (e.g. ammonium, nitrate or nitrite) is limiting in some ecosystems, such as marine systems, where this is a limiting nutrient for phytoplankton. Anaerobic ammonium oxidation (anammox) bacteria convert ammonium and nitrite into dinitrogen gas, which, along with denitrification, is one of the two pathways to remove bioavailable nitrogen. Marine anammox is mediated by '*Candidatus Scalindua*' bacteria and is responsible for ca. 30 % of the bioavailable nitrogen lost from modern oceans (1). Tracking these microorganisms is therefore essential to model the modern and paleo nitrogen cycle. Lipid biomarkers can be used to identify anammox processes in the geological nitrogen cycle. However, the diagnostic anammox biomarker of ladderane lipids are highly strained cyclobutyl moieties and are not recalcitrant in sediments older than 125,000 years (2). Bacteriohopanepolyol (BHP) lipids are produced by a wide range of bacteria. They function as membrane fluidity regulators in response to environmental stress (e.g. temperature, pH). Some BHPs, such as bacteriohopane-17 $\beta$ ,21 $\beta$ (H), 22*R*, 32*R*, 33*R*, 34*S*, 35-tetrol (BHT-34*S*; Fig 1), are ubiquitous in the environment, and so are not diagnostic biomarkers. However, a BHT diastereomer with unknown stereochemistry (BHT-*x*) has only one known producer, the marine anammox bacteria '*Ca. Scalindua*'. This biomarker can persist in the environment at least 2 million years (2). Varying retention times by gas chromatography (GC) were used to distinguish BHT-*x* from BHT-34*S* and BHT-34*R* (Fig 1) (3), allowing this lipid biomarker to be applied to identify anammox bacteria in the geological record (2). Directly establishing the stereochemistry of BHT-*x* by NMR of sufficient pure lipid obtained from the bacteria is hindered by the extremely slow growth rate of '*Ca. Scalindua*'. To fully utilize this biomarker, alternative methods for structural identification are required.

To understand and in-turn create predictive models for BHPs compounds, we have applied a 'low-cost', composite DFT method, B97-3c, to calculate molecular polarizability, HOMO-LUMO gap and dipole moment. These calculations build upon preliminary work utilising the inexpensive semi-empirical PM3 method to calculate the molecular polarizability (4). The B97-3c method has been shown to yield excellent molecular and condensed geometries, similar to most hybrid functionals evaluated in a large basis set expansion (5). Therefore, we propose its use to build upon previous work and aid the assignment of unknown BHPs.

### *Methods*

All optimisations of isolated BHP molecules and quantum descriptors (molecular polarizability, FMO energies and dipole moment) were calculated at the B97-3c level of theory (5), with Orca

4.2.1 program (6) and using Gabedit 2.5 (6) for visualisation. Relative binding energies of BHP molecules to the column was modeled using the semi-empirical PM3 method. A short oligomer of poly-dimethylsiloxane was chosen to represent the column structure, with its length extending beyond that of the BHP molecule. Due to the size of the BHP column adduct full DFT calculations were not possible in this project. The gas phase energy difference between the BHP adduct and the isolated species were used to calculate the relative binding energies.

GC retention times for all BHPs were compiled from previous and novel analysis (8), analysed using two different columns. One column was selected for full separation of BHP diastereomers such as BHT-34S, BHT-34R and BHT-x (isomer column) and one for elution of aminoBHPs (amino column) such as aminotriol and aminotetrol (8).

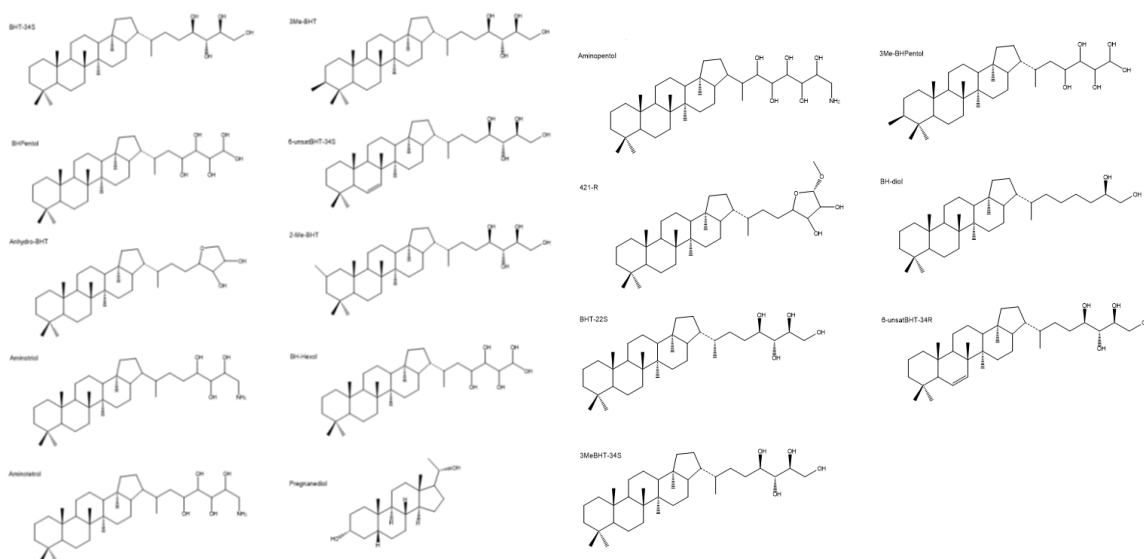


Figure 1. Structures of selected BHP biomarkers.

## Results

Interpretation of data was performed analogous to that of Allcock *et al.* (3) A line of best fit was applied to assess the correlation of retention time and polarizability on each column. The isomer column produced a positive correlation (Fig. 2), while the amino column failed to show a correlation of any significance ( $R^2 = 0.17$ ). This strong positive correlation ( $R^2 = 0.97$ ) is greater than what has been previously reported on these large, computationally intensive molecules, using a less complex PM3 method ( $R^2 = 0.89$ ). Furthermore, this modelling placed all structural groups on a single line of best fit. This further demonstrates the ability of a polarizability based approach to distinguish structural variation and complex 3-dimensional organic compounds (4). Of the 3 computational descriptors investigated using the B97-3c method, only isotropic polarizability succeeded in producing a positive correlation. Binding energy calculations were also carried out using the less computationally intensive PM3 method, but these failed to produce a significant correlation.

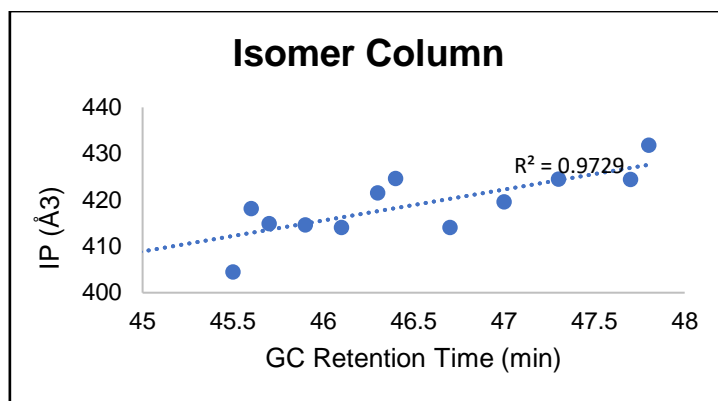


Figure 2. Correlation of polarizability and retention time of the isomer column.

### Discussion

A positive linear correlation was found between polarizability and GC retention times for structures of varying size and/or stereochemistry (Fig. 2) with the B97-3c method. A significant improvement in the data fit compared to previous semi-empirical methods was observed. GC retention times are primarily controlled by volatility. Both molecular boiling point and polarizability increase with molecular size, corresponding to a positive relationship with GC retention time. However, the polarizability increases of BHPs with varying stereochemistry (e.g. BHT-34S vs. BHT-34R; 2Me-BHT-34S vs. 3Me-BHT-34S) are generally consistent with increasing GC retention time (Fig. 2). This suggests that the correlation of polarizability to GC retention time is caused by additional factors to boiling point increases with increasing molecular size.

### Conclusion

BHP polarizability and GC retention time are positively correlated, further demonstrating that polarizability is a controlling factor on BHP retention time. The B97-3c scheme shows promise in the application for *in silico* identification of unknown BHP stereochemistry. Future work will apply the B97-3c method to the binding energy calculations, to develop a correlation between this molecular descriptor and retention time.

### Acknowledgments

Funding was obtained through the Chromatographic Society (ChromSoc) and British Mass Spectrometry Society (BMSS) awarded to RSN, TN, and AH.

- [1] B.B. Ward. *Science* 2013, 341, 352-353
- [2] D. Rush, H. M. Talbot, M. T. J. van der Meer, E. C. Hopmans, B. Douglas, J. S. Sinninghe Damsté. *Biogeosciences* 2019, 16, 2467-2479
- [3] R. Schwartz-Narbonne, P. Schaeffer, E. C. Hopmans, M. Schenese, E. A. Charlton, D. M. Jones, J. S. Sinninghe Damsté, M. F. Ul Haque, M. S. M. Jetten, S. K. Lengger, J. C. Murrell, P. Normand, G. H. L. Nuijten, H. M. Talbot, D. Rush. *Org. Geochem.* 2020, 143, 103994
- [4] Libby Allcock, Scott Hardy, Claire Vassallo, Alex Hamilton, Tim Nichol, Rachel Schwartz-Narbonne, *41st BMSS Annual Meeting 2021*
- [5] T. Yokoyama, S. Ohuchi, T. Matsui, Y. Kaneko and T. Sasagawa, *J. Phys. Chem. C*, 2021, 125, 6601-6610
- [6] F. Neese, *WIREs Comput. Mol. Sci.* 2018, 8:e1327. doi: 10.1002/wcms.1327
- [7] A. R. Allouche, *J. Comput. Chem.*, 2011, 32, 174-182.
- [8] R. Schwartz-Narbonne, J. Blewett, S.K. Lengger, P. Schaeffer, R. Pancost, D. Rush. *29th International Meeting on Organic Geochemistry 2019*