Electronic Supplementary Material (ESI) for RSC Advances

Ionic Liquid-Based Optoelectronic Sensor Arrays for Chemical Detection

Electronic Supplementary Information

Waduge Indika S. Galpothdeniya[1], Kevin S. McCarter[2], Sergio L. De Rooy[1], Bishnu P. Regmi[1], Susmita Das[1], Farhana Hasan[1], Attres Tagge[1], and Isiah M. Warner*[1]

[1]Department of Chemistry, Louisiana State University, Baton Rouge, LA70803, USA

[2]Department of Experimental Statistics, College of Agriculture, Louisiana State University, Baton Rouge, LA70803, USA

* Corresponding author, Isiah M. Warner

Email: iwarner@lsu.edu, Phone: 1-225-578-2829

Approaches to quantifying a model's predictive accuracy

One approach is to use the model to classify the observations that were used in constructing the model and calculate the proportions of observations correctly and incorrectly classified. In the SAS documentation and output this is called the resubstitution method, and the resulting estimate of the error rate is called the apparent error rate. This method is easily implemented, but it tends to overstate the predictive ability of the model.[1] That is, the estimated error rates tend to be too low and the estimated accuracy rates too high. This stems from the fact that the data being classified were used to construct the model. The model therefore tends to do a better job of classification for that dataset than it would for other datasets in general. An approach that gets around this problem is the cross-validation method. Using cross-validation, each observation in the dataset is classified using a model that is constructed from a dataset that excludes the observation to be classified. This method is more complicated to implement, but has the advantage of providing more reasonable, less biased estimates of the correct and incorrect classification rates of a model.[1] Finally, predictive models can be updated as more information becomes available. To improve a model's accuracy and reliability, it can be updated using additional data for chemical substances that are already in the model's substance database. To broaden a sensor array's applicability, the predictive model's chemical substances database can be expanded by updating the model with data for new substances.

The rationales for considering the two approaches in constructing the discriminant models

The discriminant model uses estimates of certain underlying distributional characteristics in its calculations. The two main categories of parametric discriminant models being used here are linear and quadratic discriminant models. Quadratic discriminant models are more generally appropriate, but require more data to fit. Linear discriminant models require less data to fit, but are appropriate in less general circumstances. The question of whether a linear discriminant model is adequate in a given situation is a statistical question that is answered in the context of the discriminant analysis model-building process. The problem is that if there is not enough data to fit a quadratic discriminant model, there will not be enough data to perform the statistical analysis to answer the question of whether a linear discriminant model is adequate. In some cases, there may not even be enough observations to fit a linear discriminant model. In situations where there is enough data to a linear discriminant model but not enough to fit a quadratic model, one approach is to fit a linear discriminant model, and to issue the disclaimer that there is not enough data to formally test whether a quadratic model would be better.

Another way to get through the problem is to reduce the dimensionality of the predictor space, either by using only a subset of the measured predictor variables, or by constructing a small number of new variables based on the original measured variables. When the number of predictor variables is small enough relative to the number of observations within the each classification group, the process alluded to above for choosing between quadratic and linear discriminant models can be performed. Choosing a subset of variables to use can be problematic. As an example in the current context, some variables may be important for identifying certain analytes, while other variables may be important for identifying other analytes. Alternatively, PCA can be used to reduce the dimensionality of the predictor variable space while at the same time using all of the original variables.[2] If a small number of principal

components account for a sufficiently large proportion of the variability in the original variables, and if the various groups are sufficiently separated with respect to those principal components, then they may be effective as predictors in a discriminant model. All statistical analyses for this paper were generated using version 9.3 of the SAS System.

Table S1: List of misclassifications by discriminant model based on three principal component, under cross-classification.

Substance	Misclassified as Substance	Number of cases
pH1	pH7	1
pH4	pH1	1
pH4	pH7	2
pH10	pH7	1
pH13	pH7	2
CH_3COOH	CF_3COOH	1
HCl	CF_3COOH	2
HCOOH	CF_3COOH	3
Diethylamine	Triethylamine	1
DMF	Pyridine	1
Methylimidazole	CF_3COOH	1
Ammonia	Diethylamine	1
Pyridine	DMF	2
Triethylamine	Ammonia	1
Triethylamine	Diethylamine	1
Total Number of Misclassifications		21

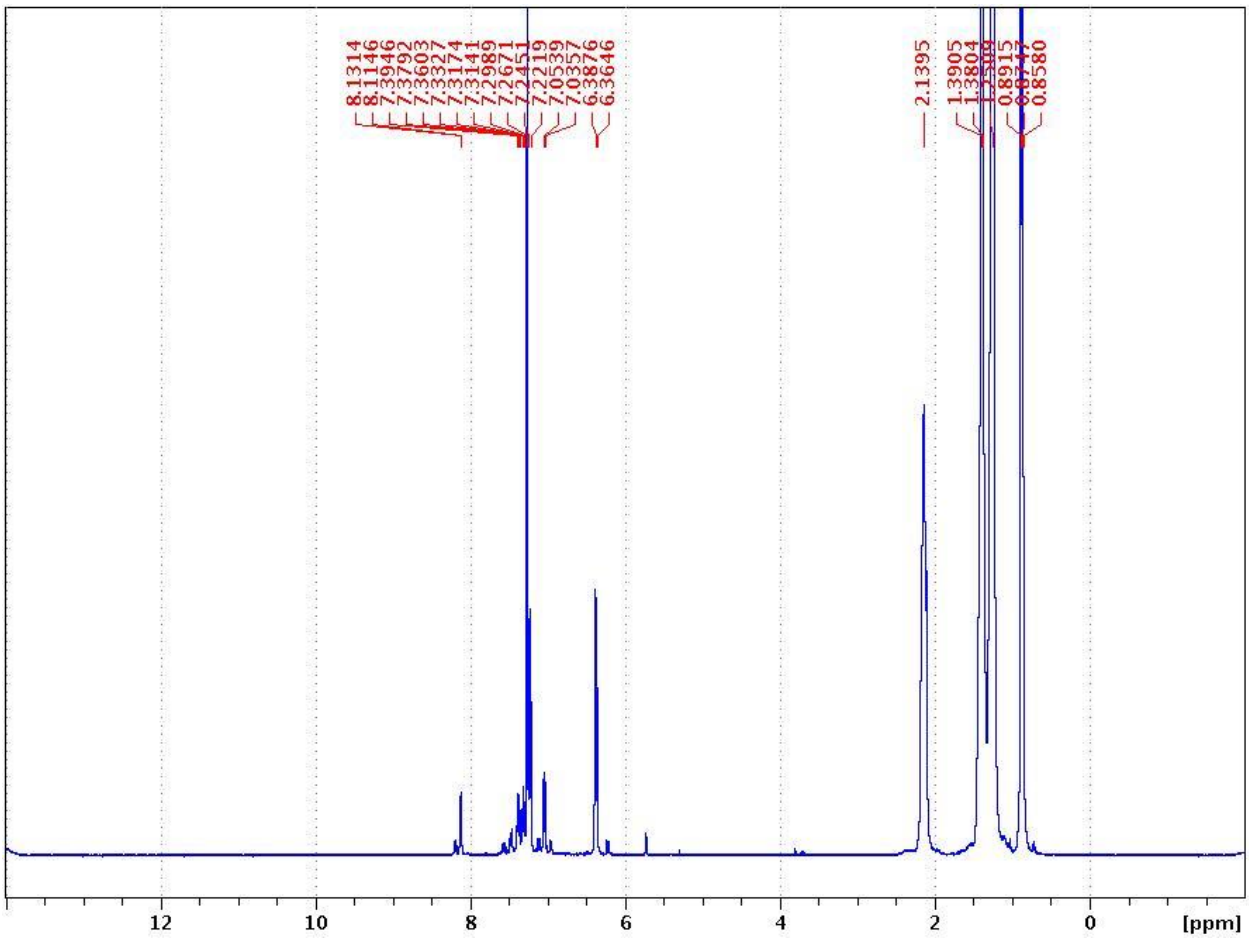


Figure S1- A. ^{1}H NMR ($CDCl_3$, 400 MHz) of P2(PR)

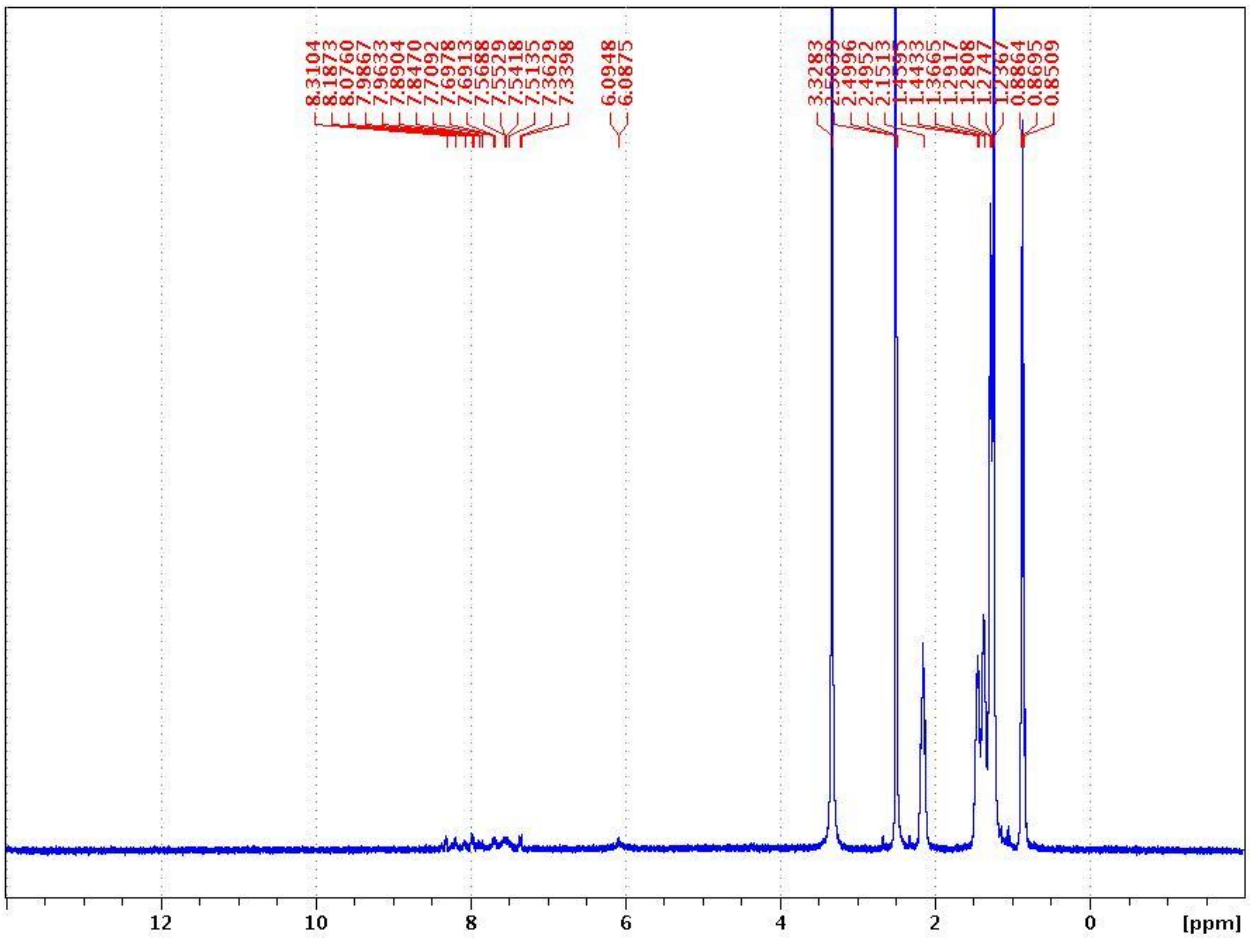


Figure S1- B. ^{1}H NMR (DMSO, 400 MHz) of P_2(BY)

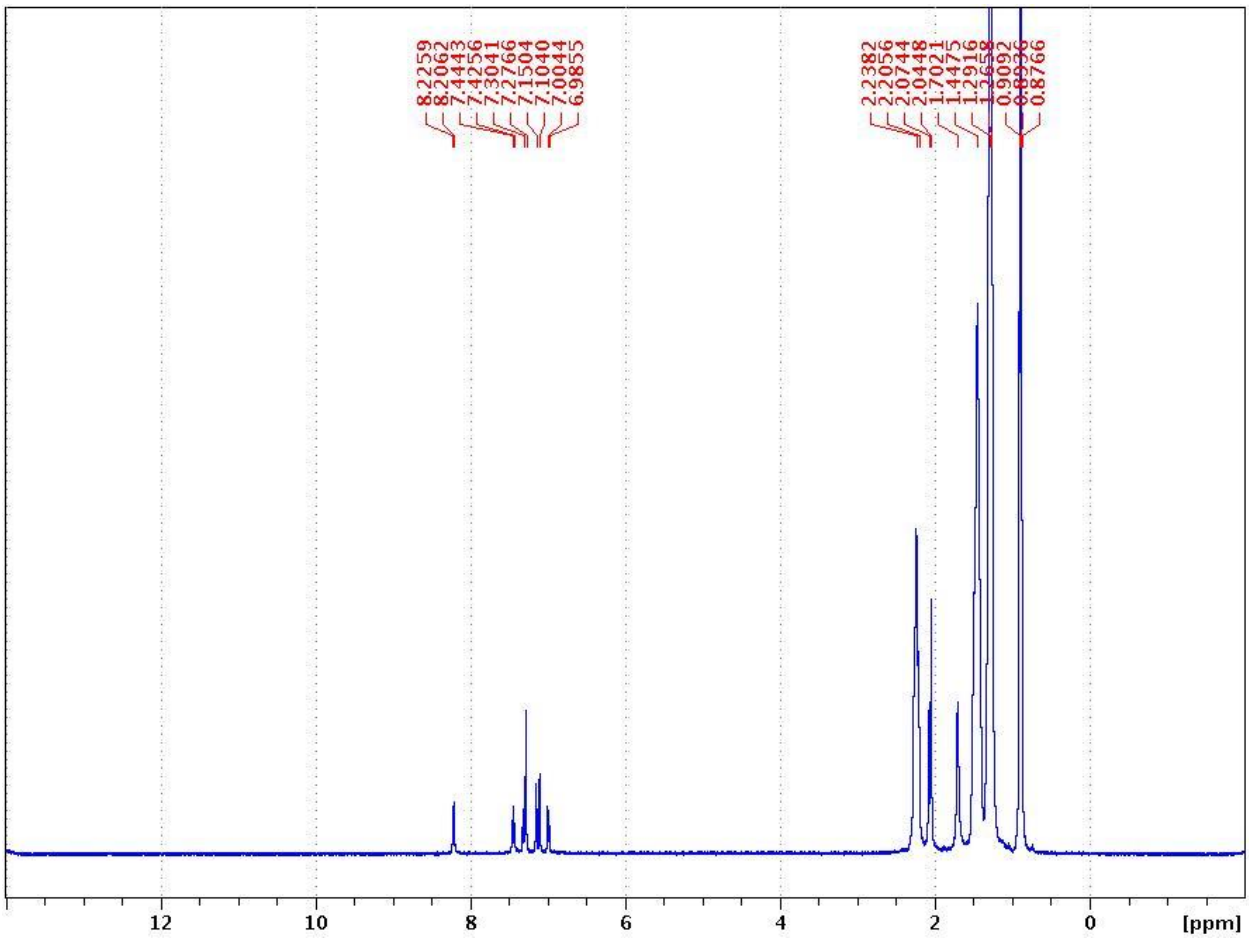


Figure S1- C. ^{1}H NMR ($CDCl_3$, 400 MHz) of P_2(BCG)

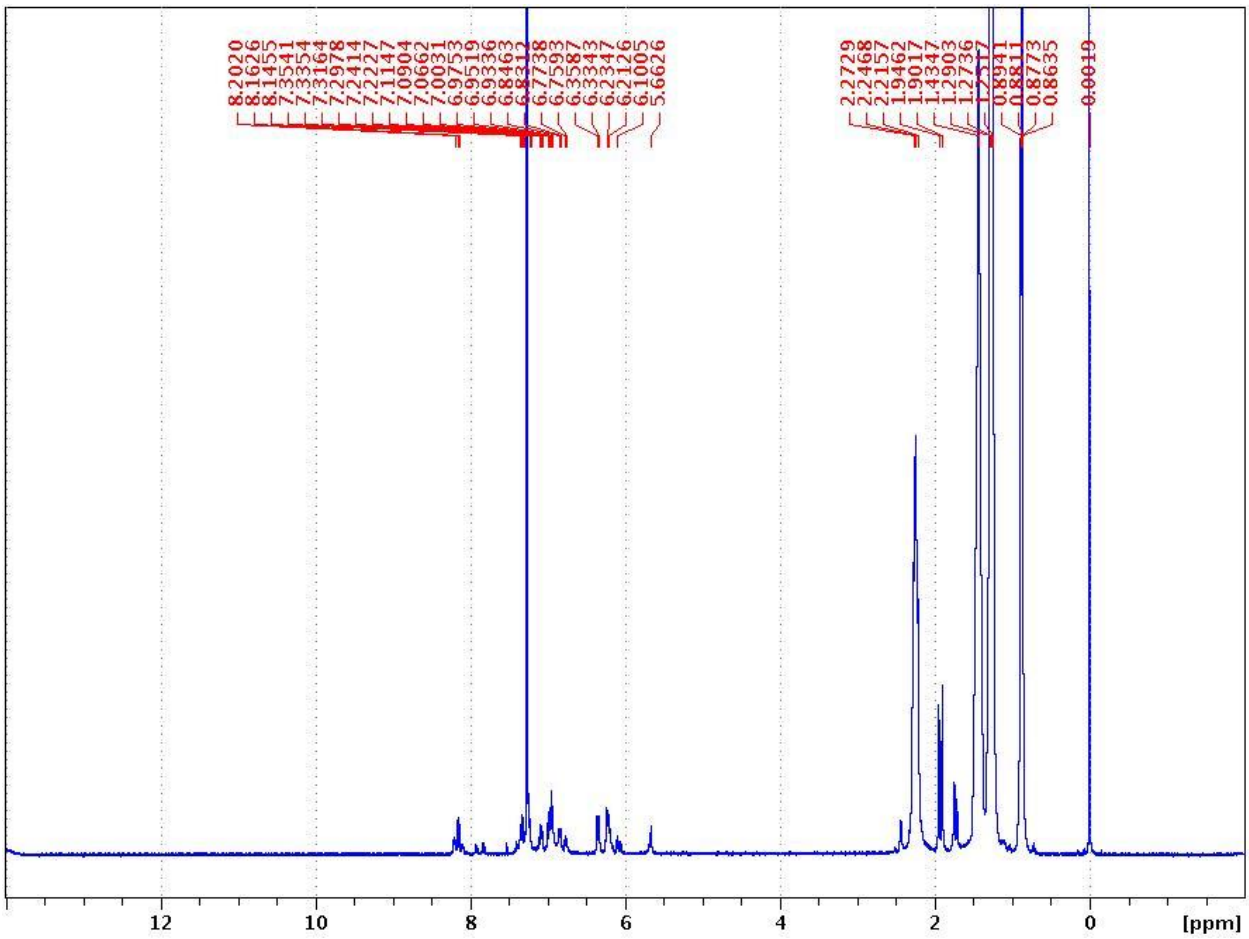


Figure S1- D. ^{1}H NMR ($CDCl_3$, 400 MHz) of P_2(mCP)

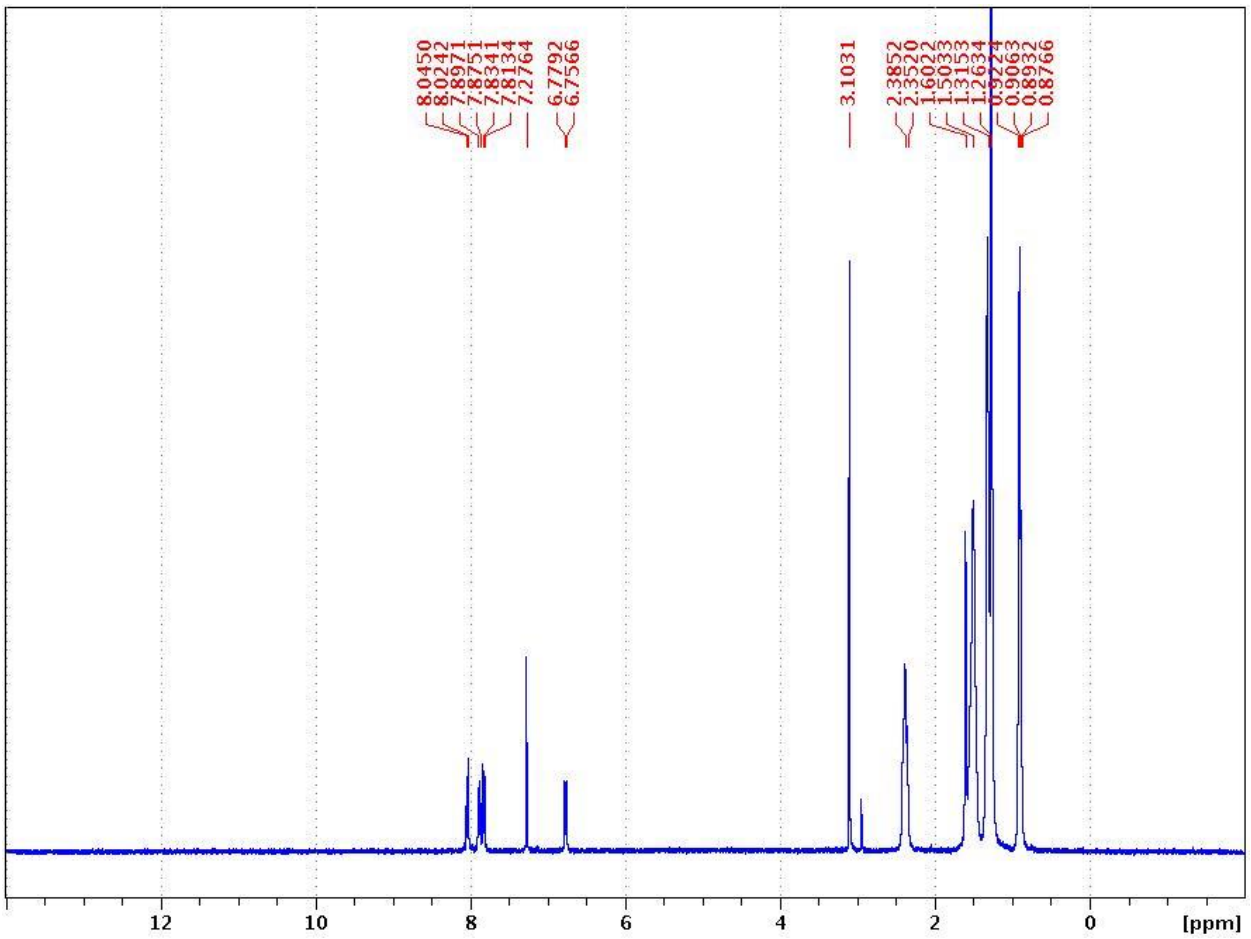


Figure S1- E. ^{1}H NMR ($CDCl_3$, 400 MHz) of P(MO)

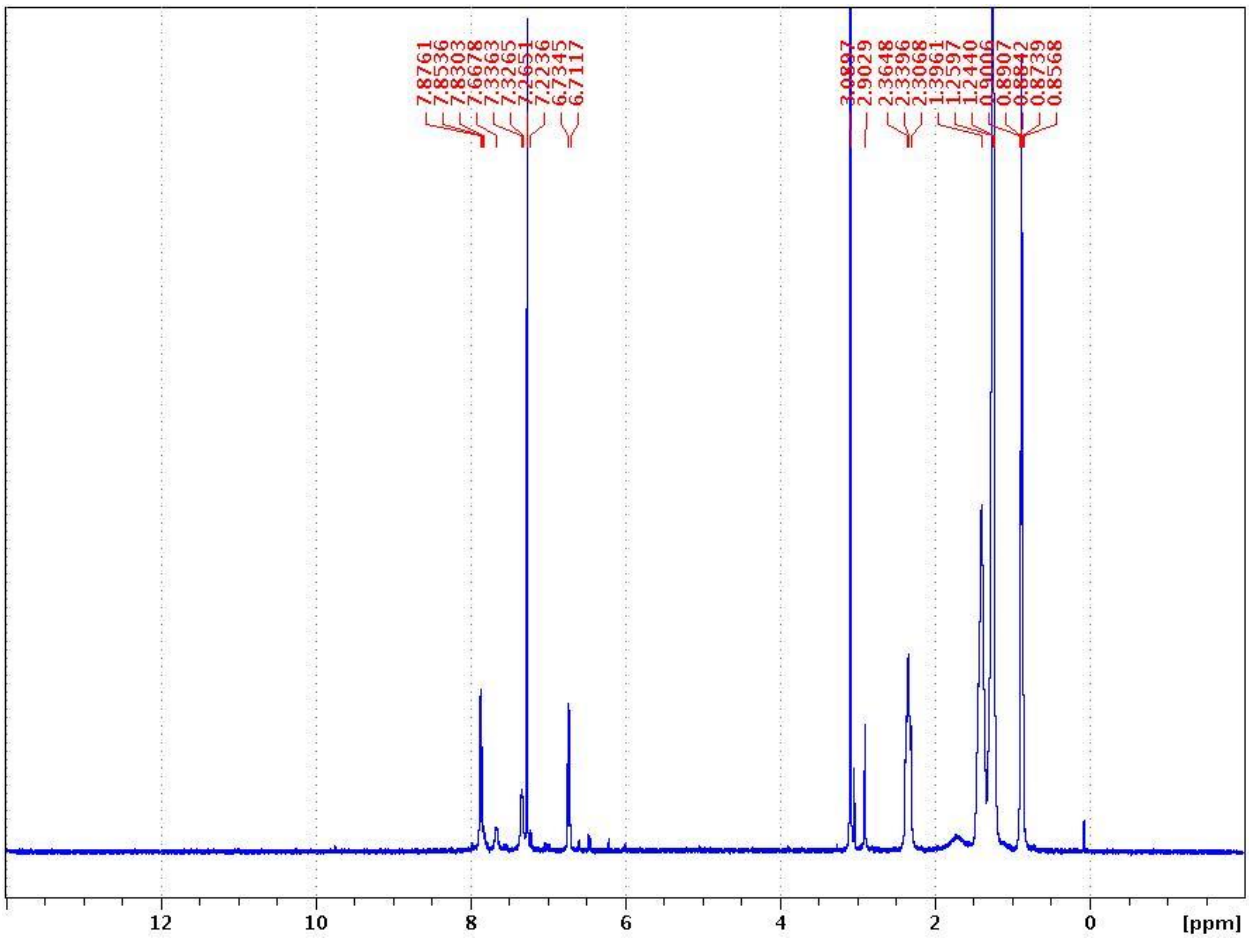


Figure S1- F. ^{1}H NMR ($CDCl_3$, 400 MHz) of P(MR)

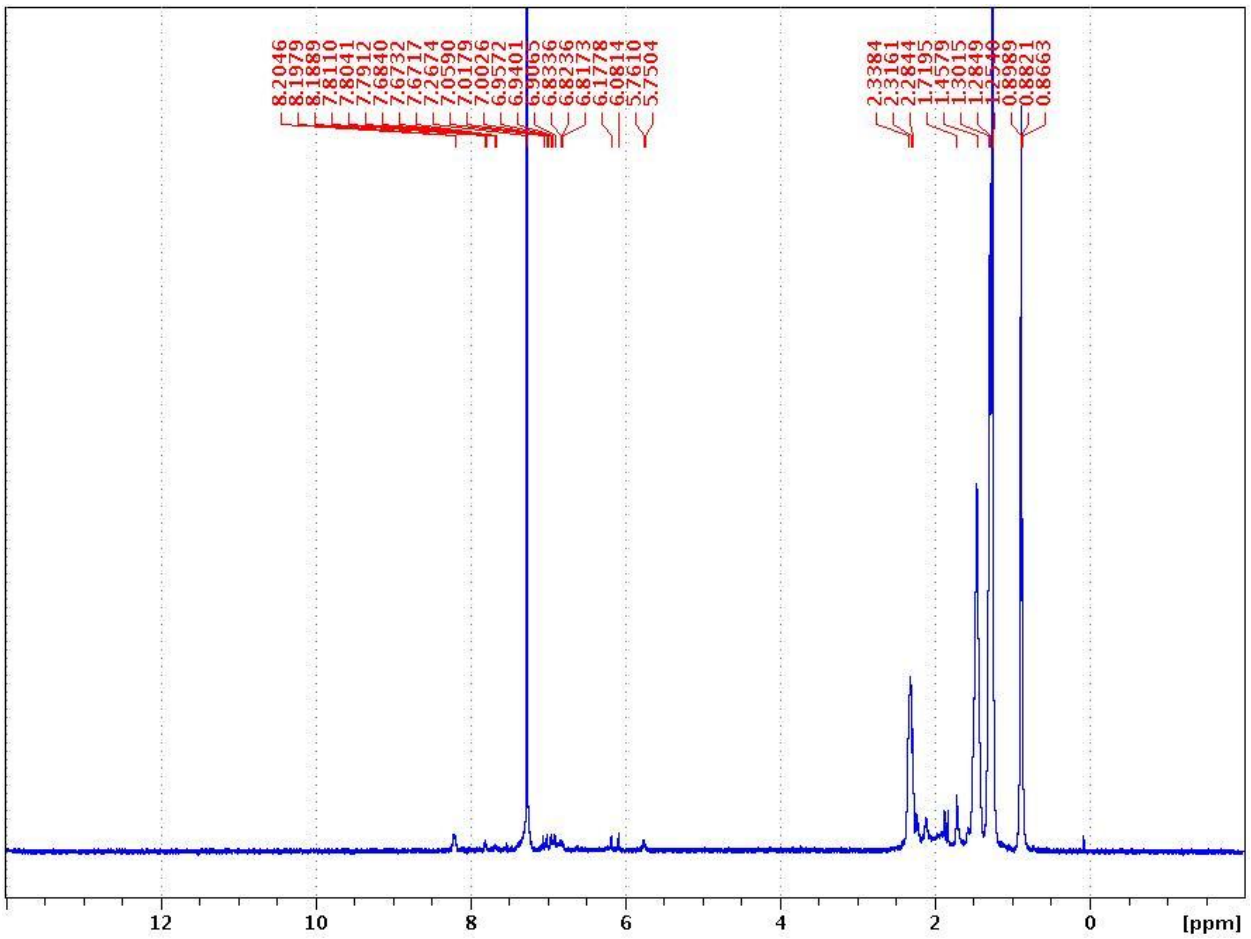


Figure S1- G. ^{1}H NMR ($CDCl_3$, 400 MHz) of P_2(Xyl)

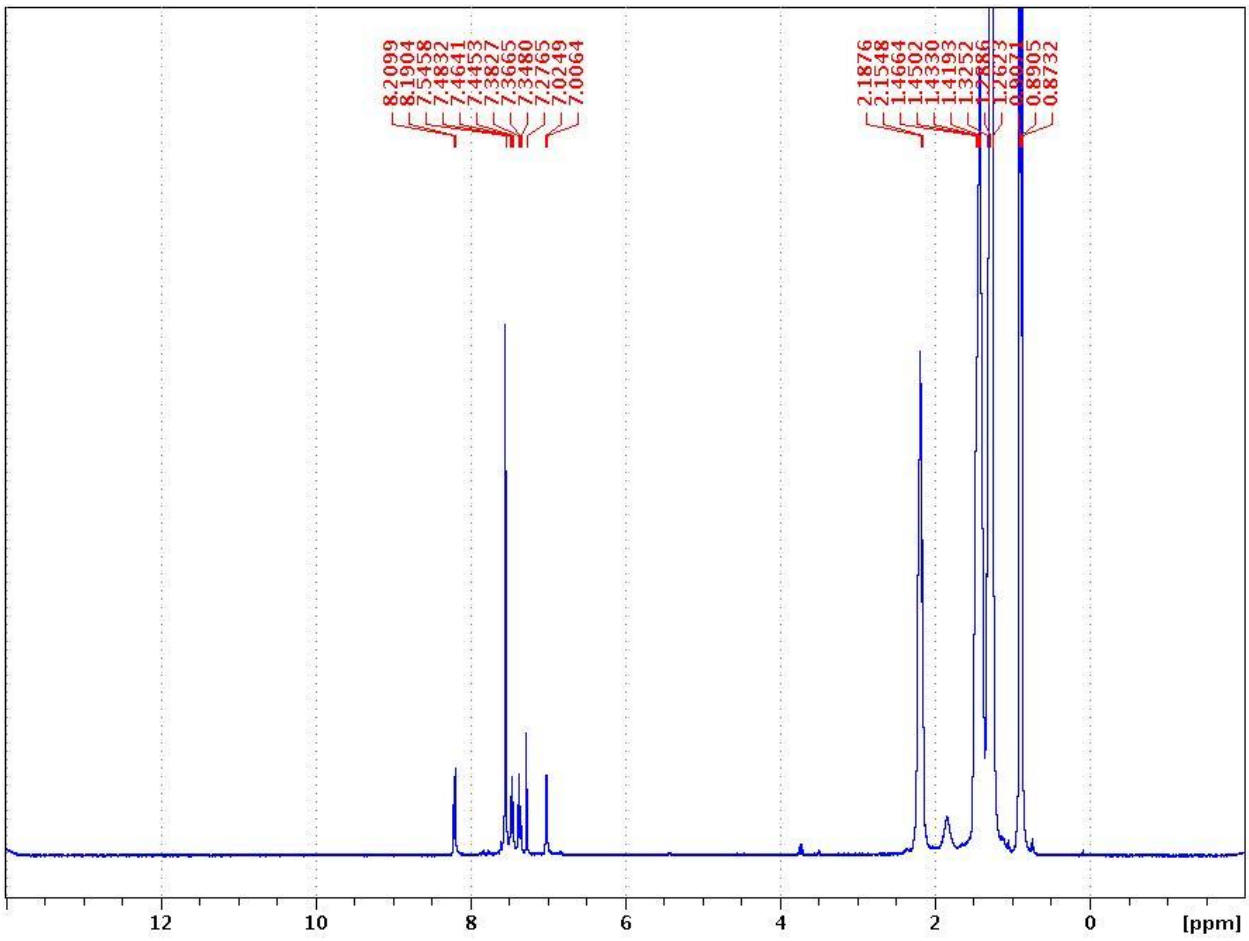


Figure S1- H. ^{1}H NMR ($CDCl_3$, 400 MHz) of P_2(BPB)

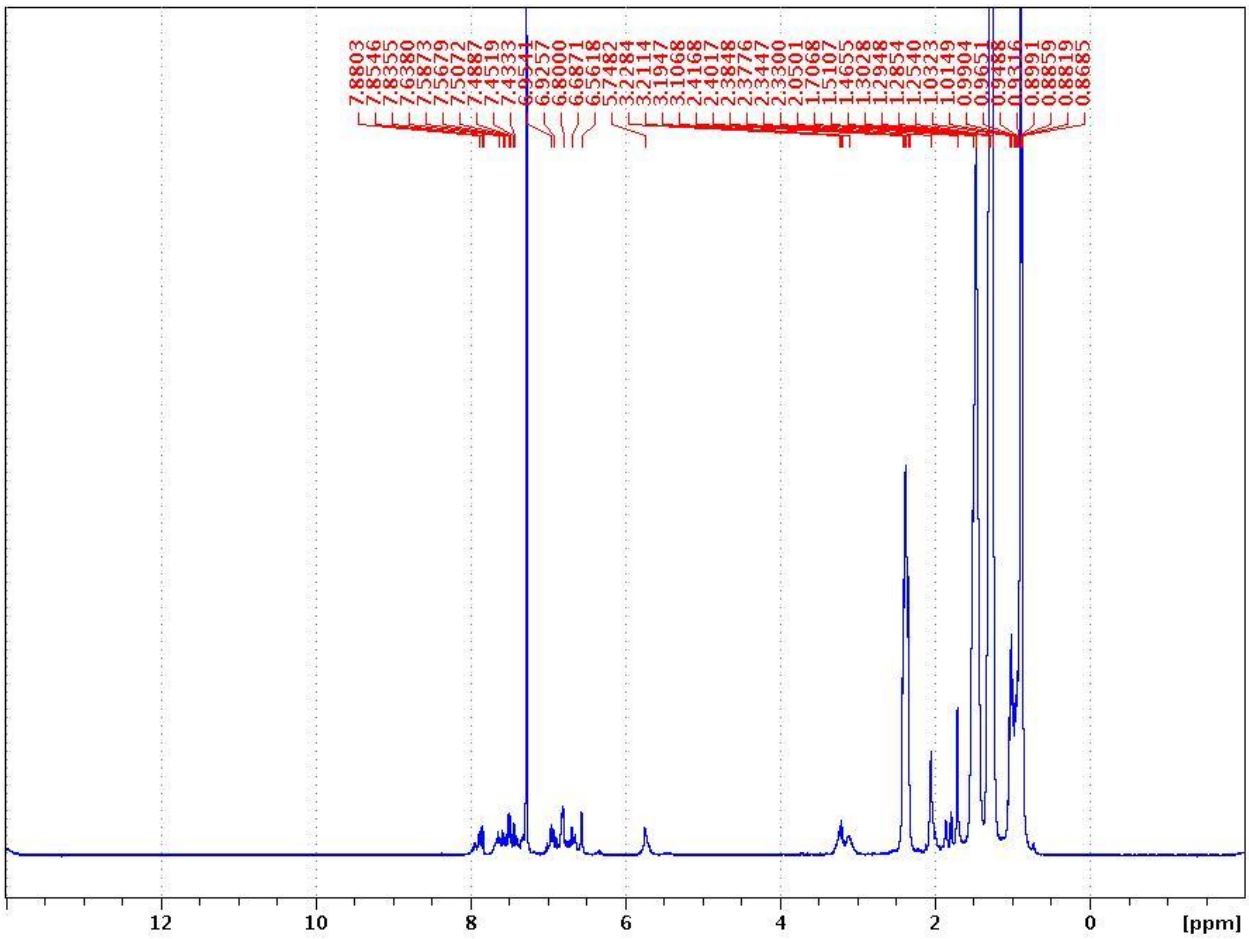


Figure S1- I. ^{1}H NMR ($CDCl_3$, 400 MHz) of P_2(Thy)

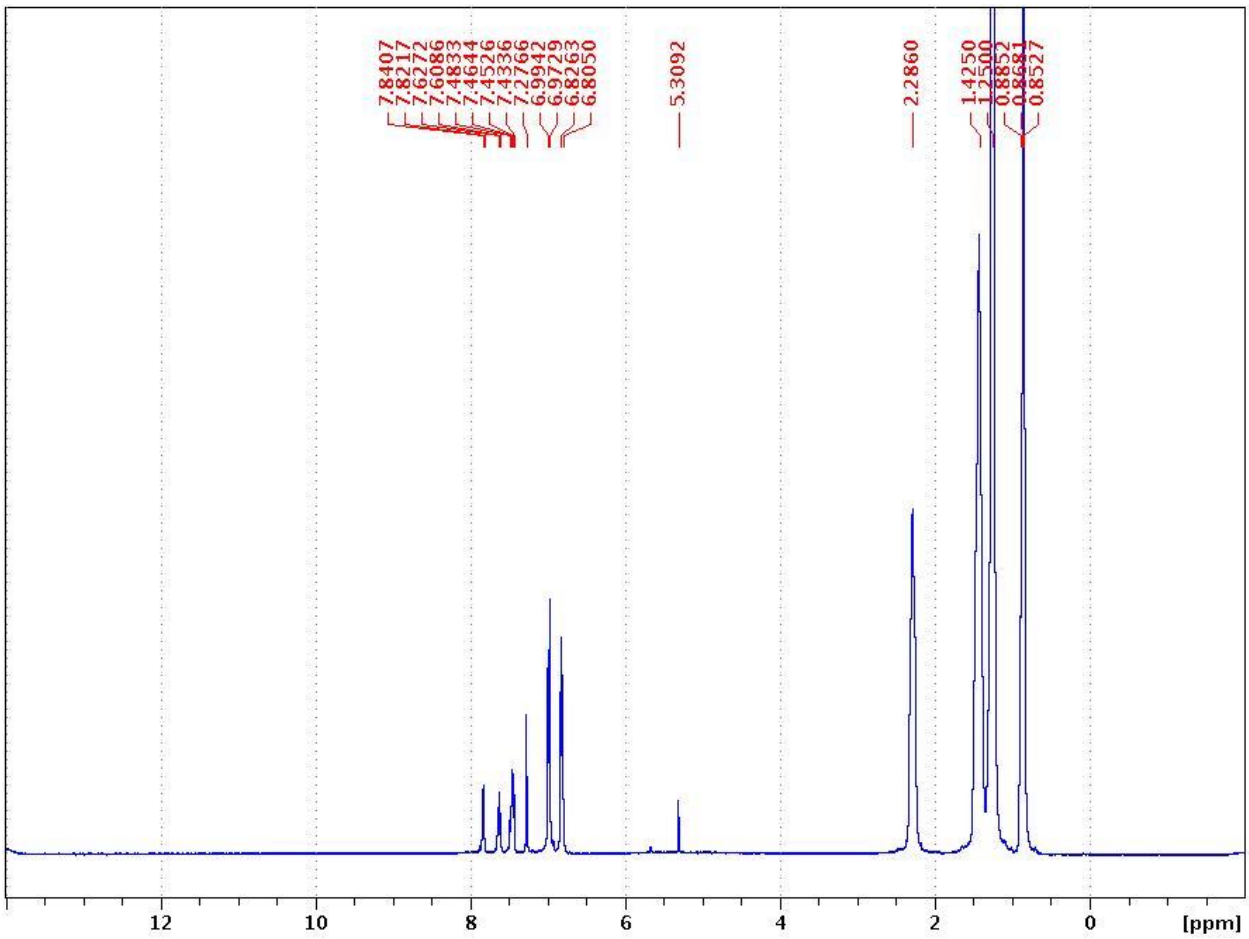


Figure S1- J. ^{1}H NMR ($CDCl_3$, 400 MHz) of P_2(FFT)

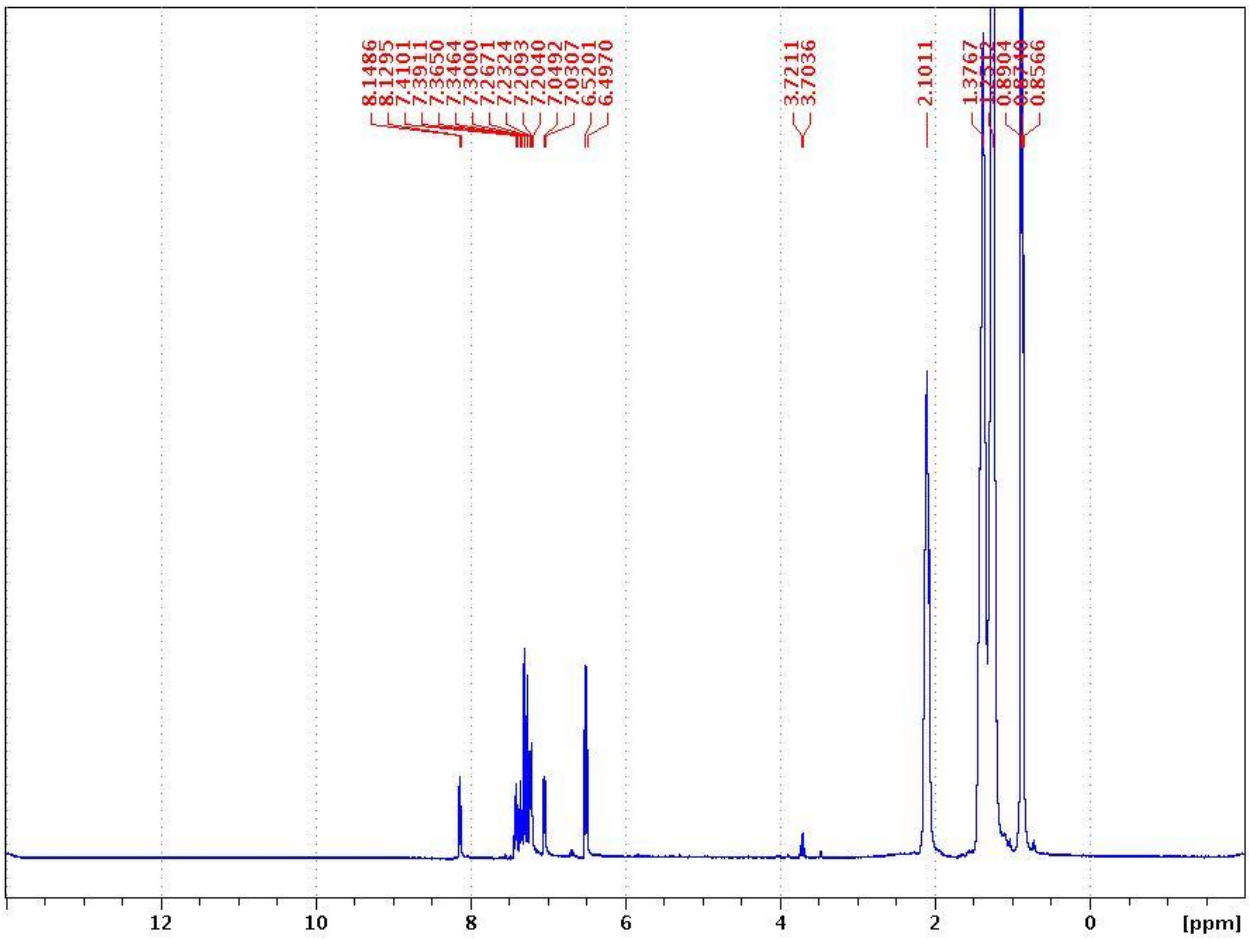


Figure S1- K. ^{1}H NMR ($CDCl_3$, 400 MHz) of P_2(ClR)

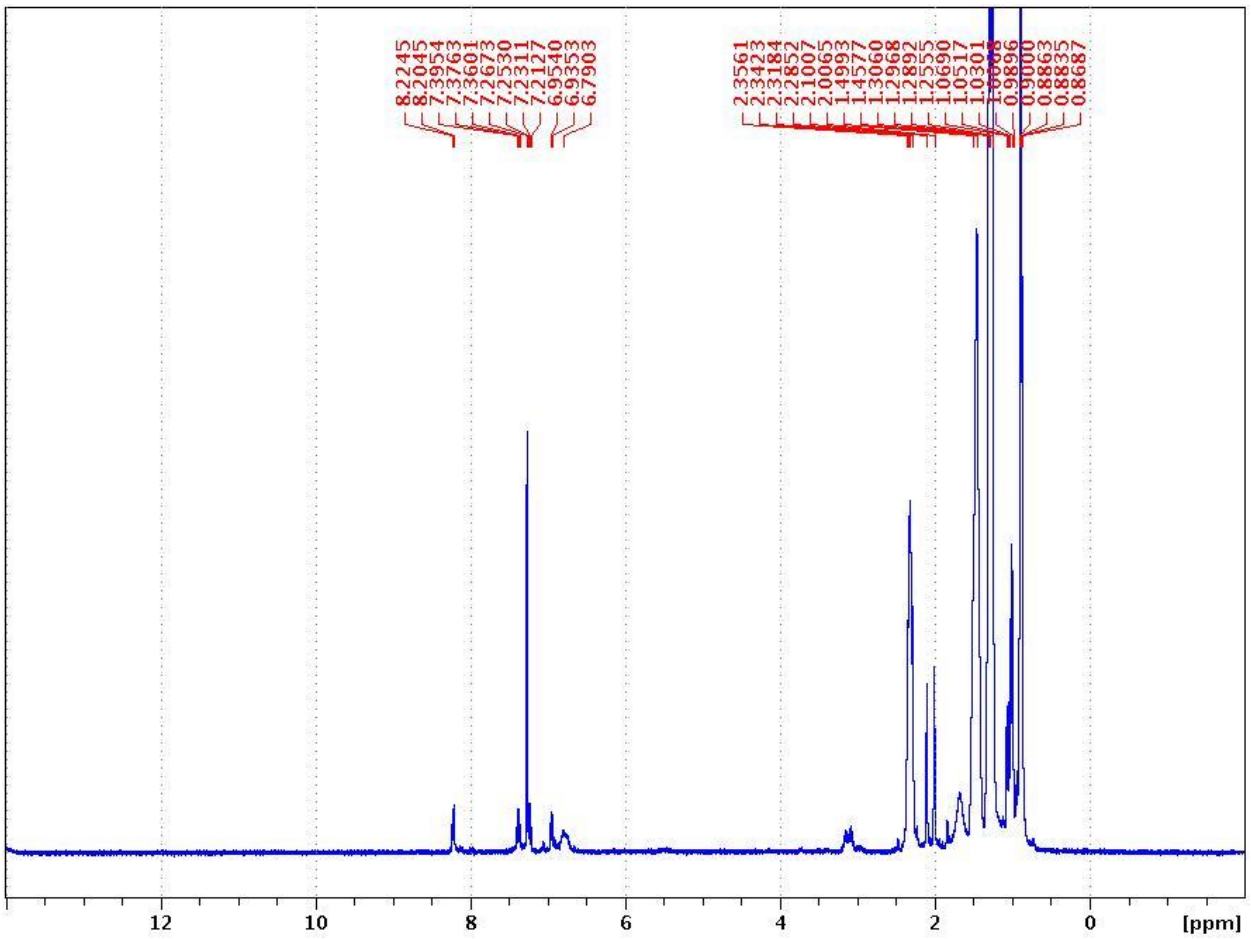


Figure S1- L. ^{1}H NMR ($CDCl_3$, 400 MHz) of P_2(BTB)

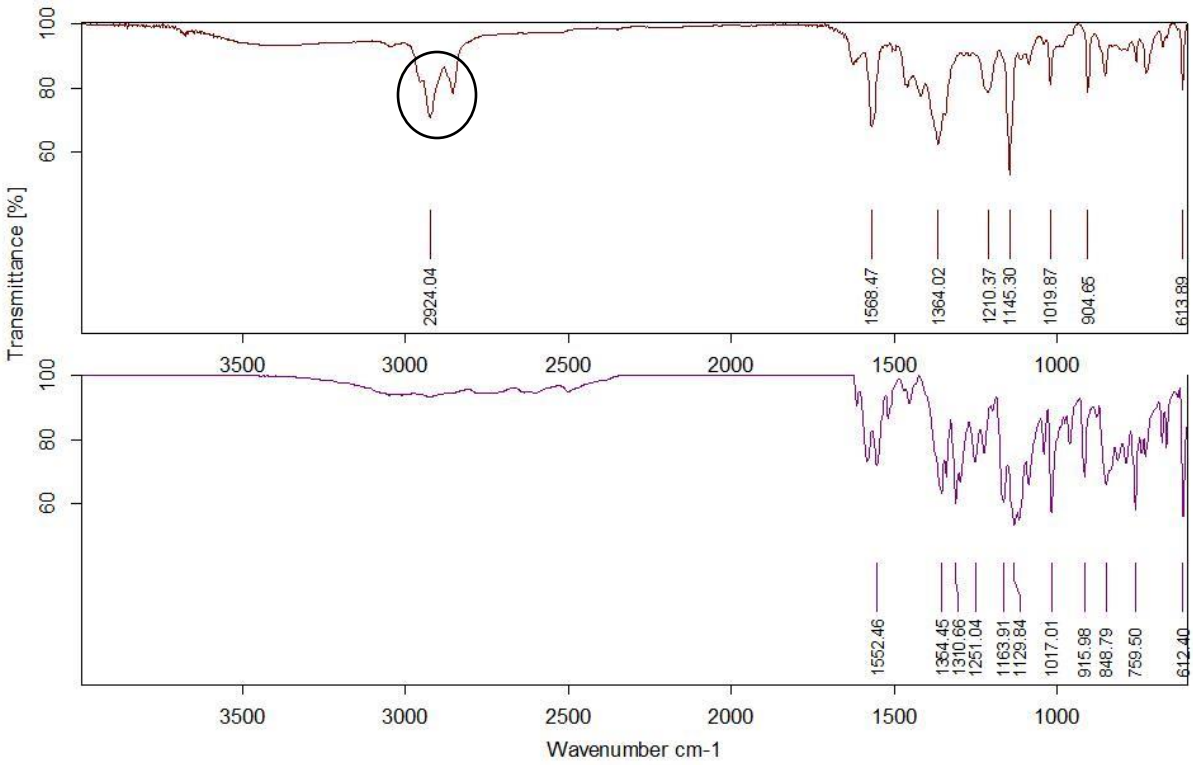


Figure S2-1A. FTIR of P_2(PR) (top) and PR (bottom) ; Circle- Characteristic C-H stretching confirms the formation of IL

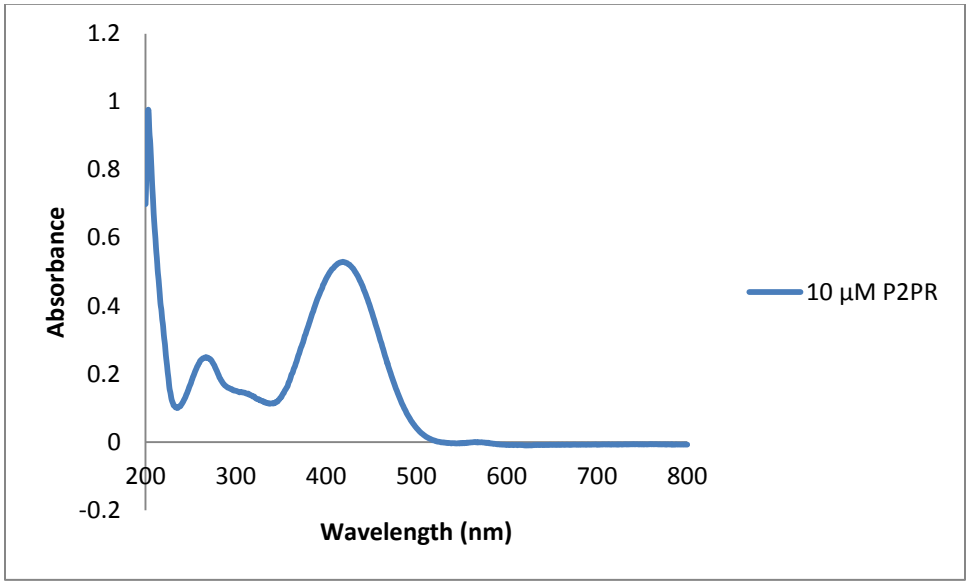


Figure S2-1B. UV-Vis spectra of P_2(PR)

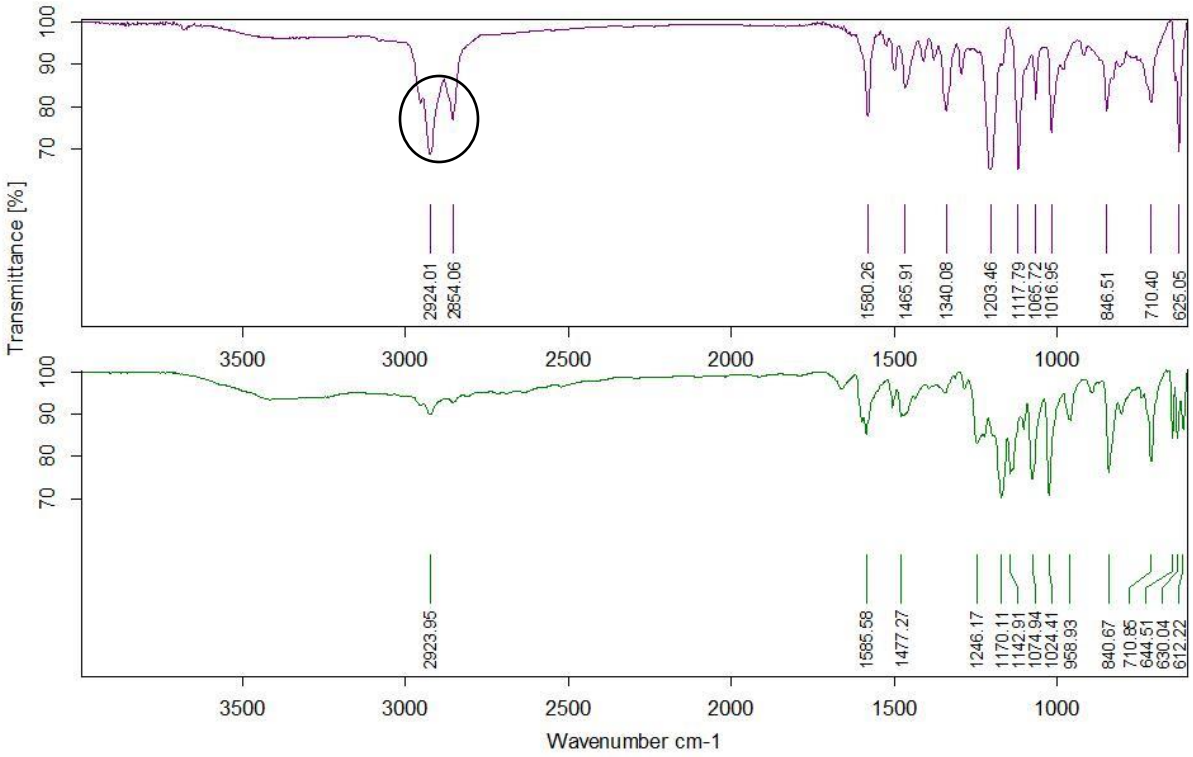


Figure S2-2A. FTIR of P_2(BY) (top) and BY (bottom) ; Circle- Characteristic C-H stretching confirms the formation of IL

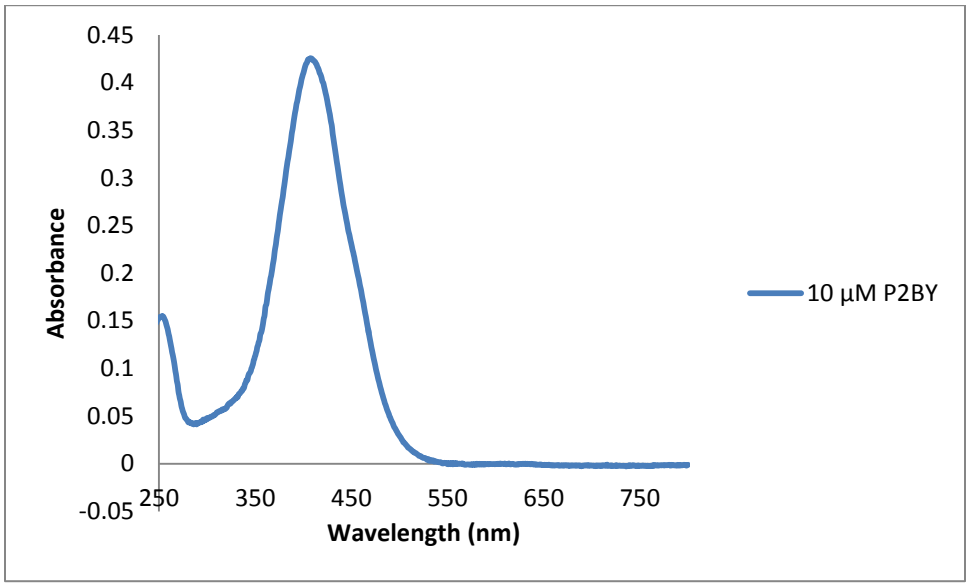


Figure S2-2B. UV-Vis spectra of P_2(BY)

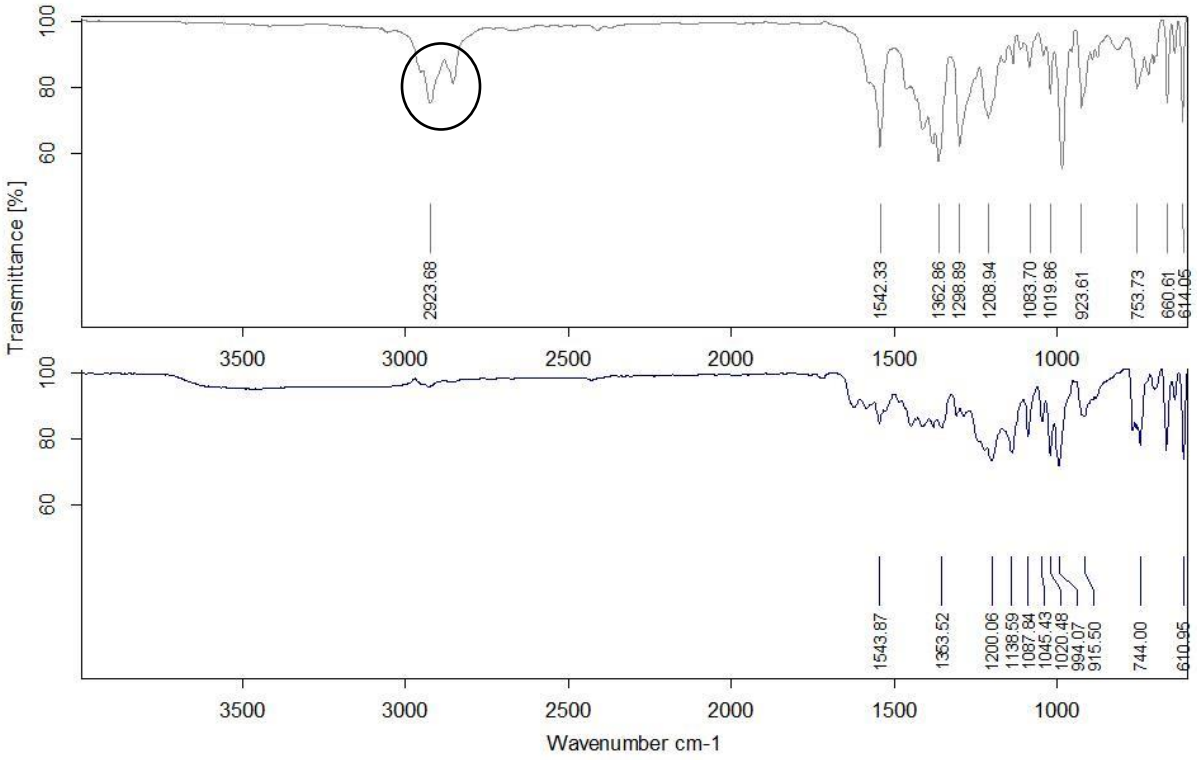


Figure S2-3A. FTIR of P_2(BCG) (top) and BCG (bottom); Circle- Characteristic C-H stretching confirms the formation of IL

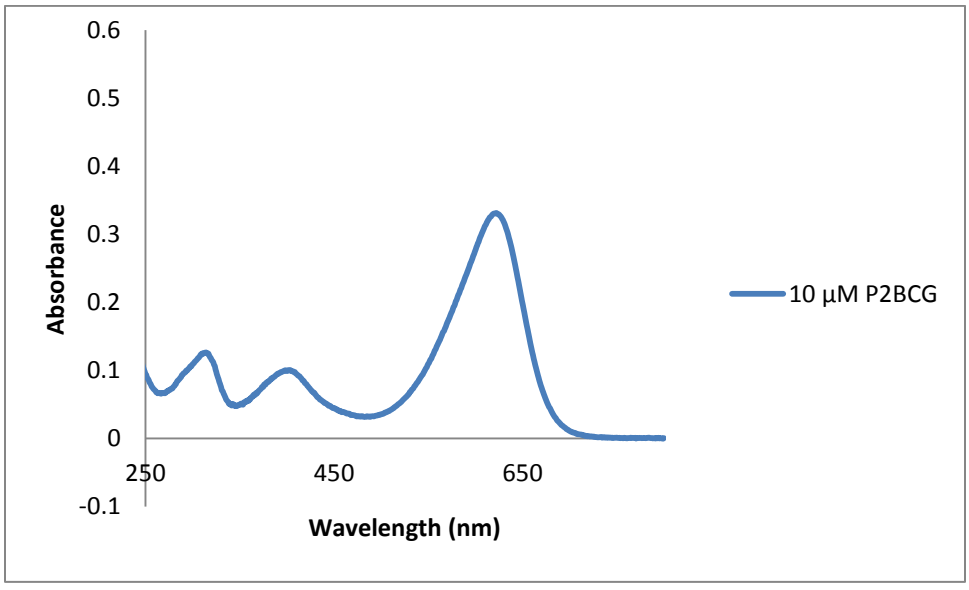


Figure S2-3B. UV-Vis spectra of P_2(BCG)

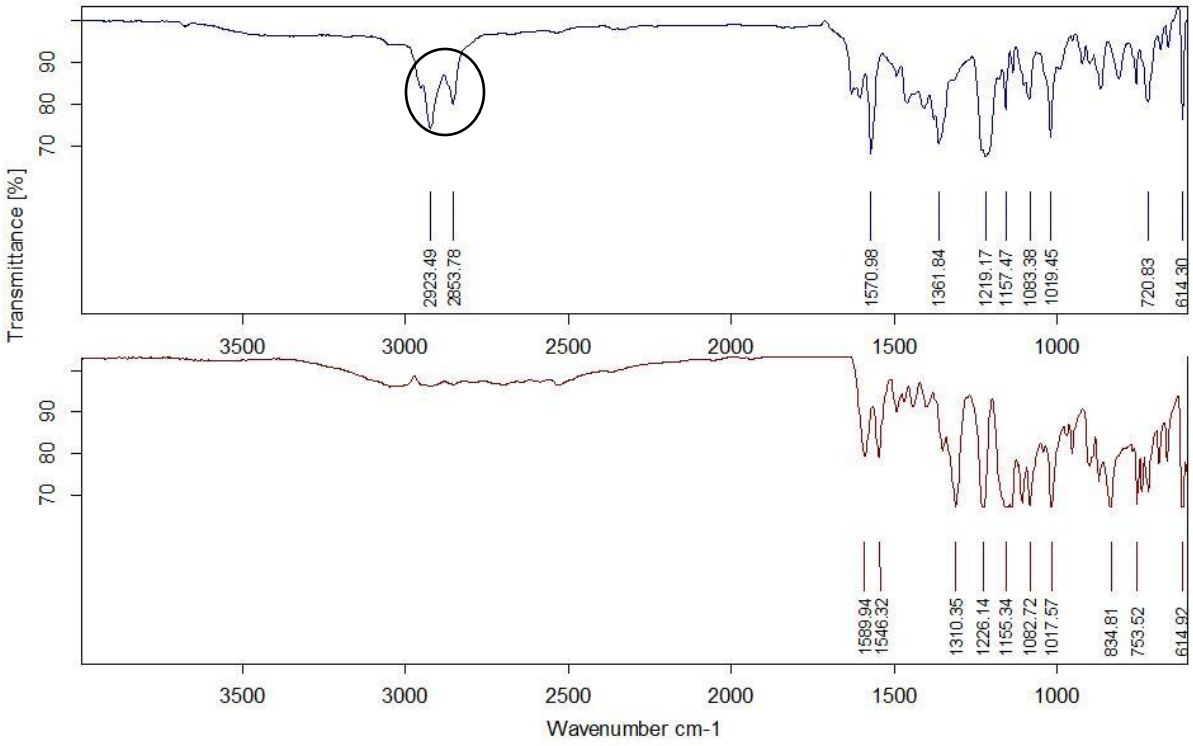


Figure S2-4A. FTIR of P_2mCP (top) and mCP (bottom); Circle- Characteristic C-H stretching confirms the formation of IL

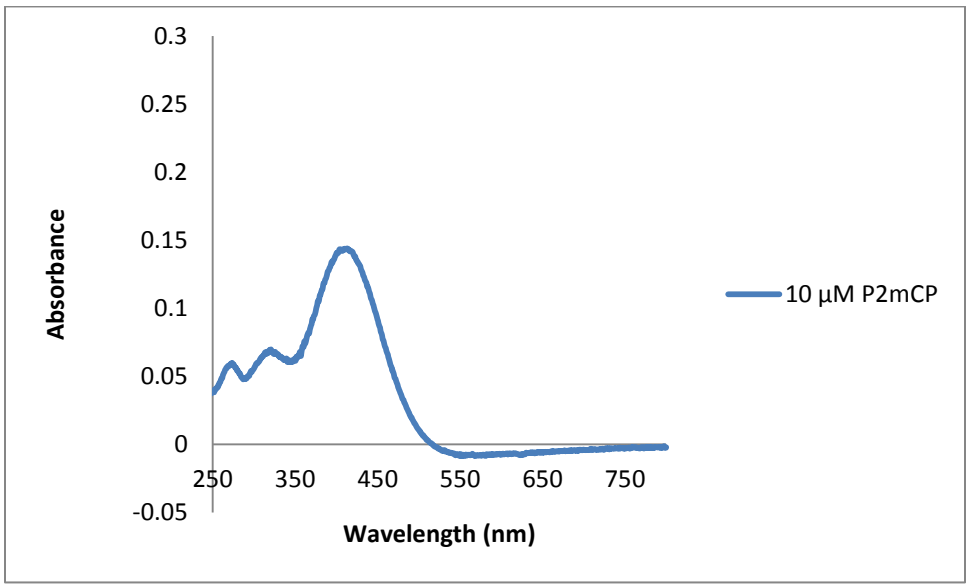


Figure S2-4B. UV-Vis spectra of P_2(mCP)

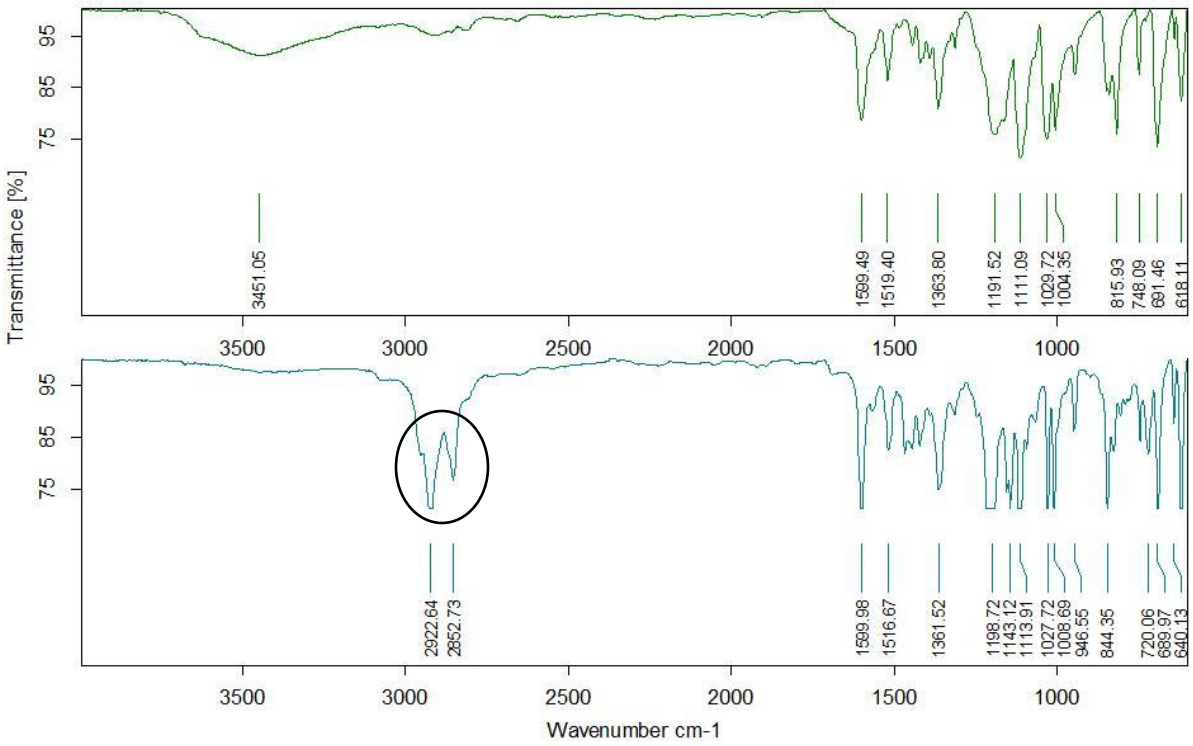


Figure S2-5A. FTIR of MO (top) and P(MO) (bottom) ; Circle- Characteristic C-H stretching confirms the formation of IL

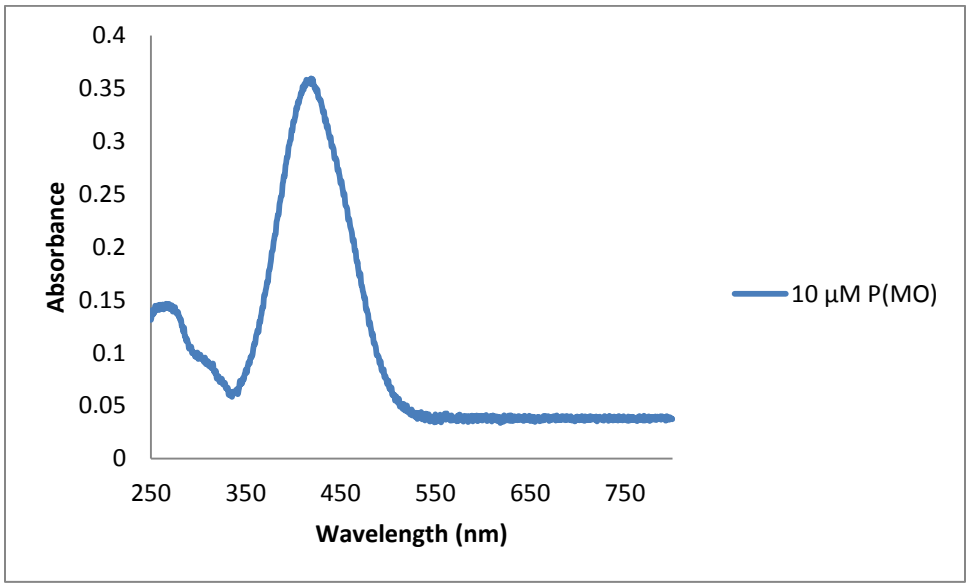


Figure S2-5B. UV-Vis spectra of P(MO)

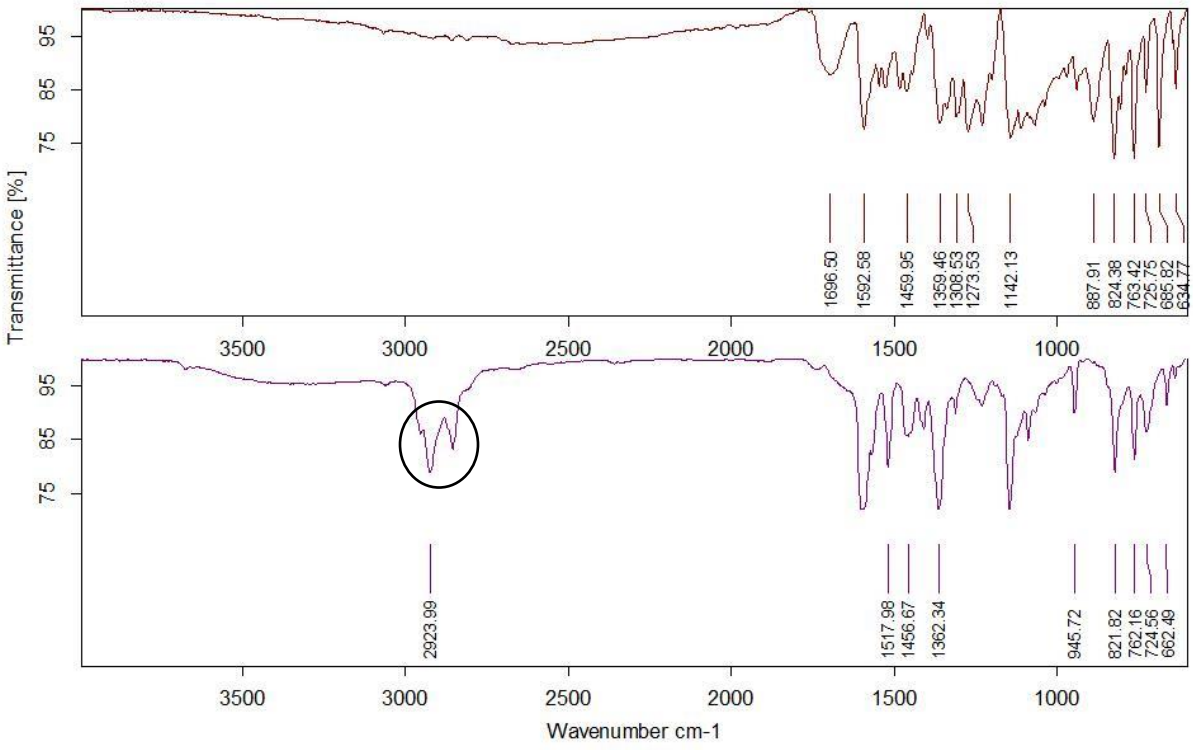


Figure S2-6A. FTIR of MR (top) and P(MR) (bottom) ; Circle- Characteristic C-H stretching confirms the formation of IL

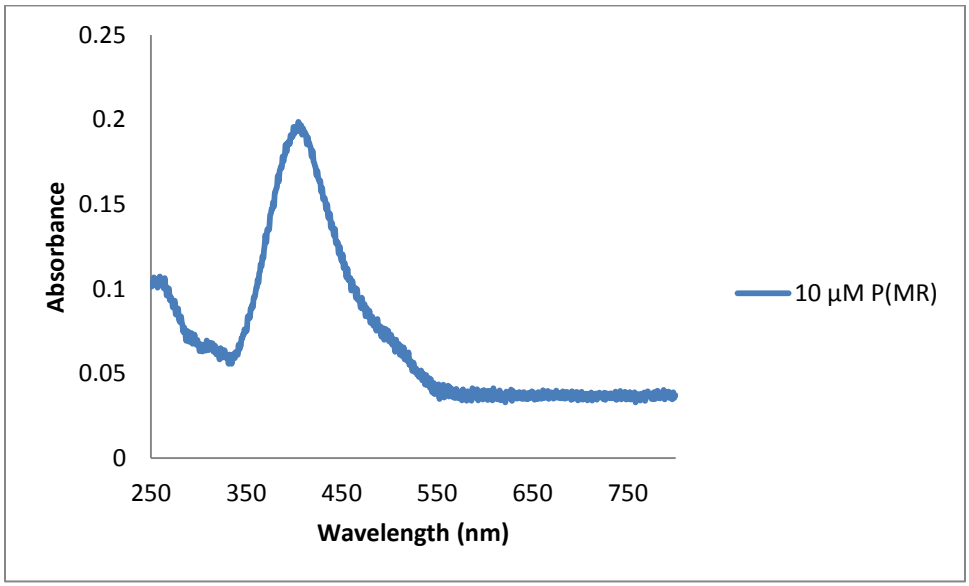


Figure S2-6B. UV-Vis spectra of P(MR)

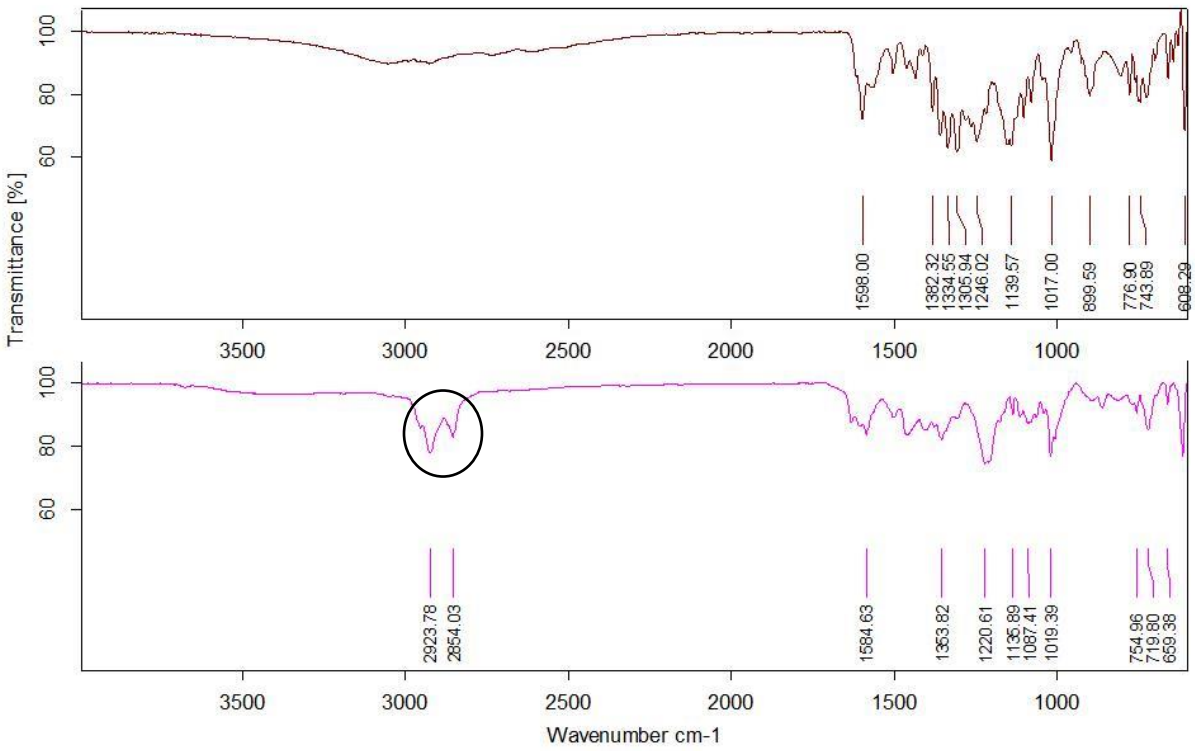


Figure S2-7A. FTIR of Xyl (top) and P_2(Xyl) (bottom) ; Circle- Characteristic C-H stretching confirms the formation of IL

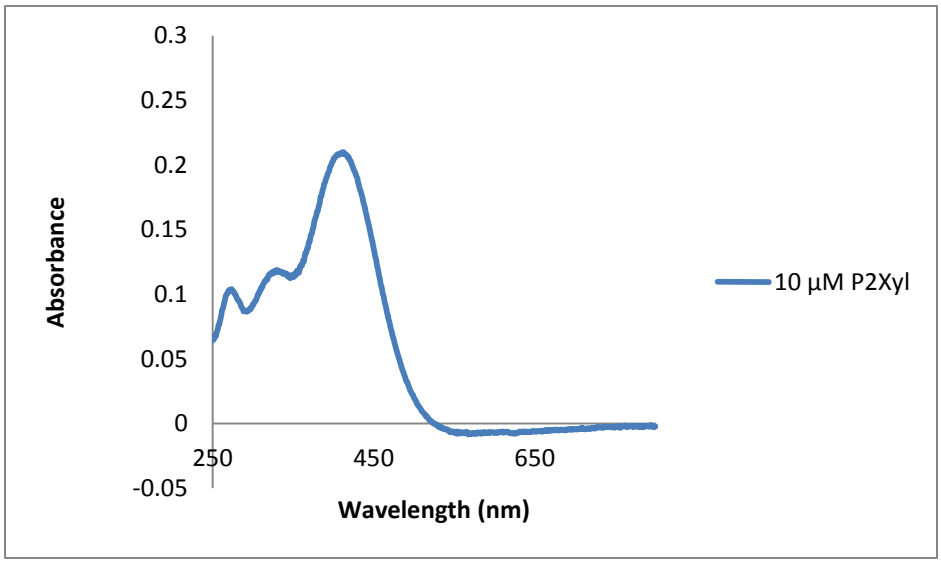


Figure S2-7B. UV-Vis spectra of P_2(Xyl)

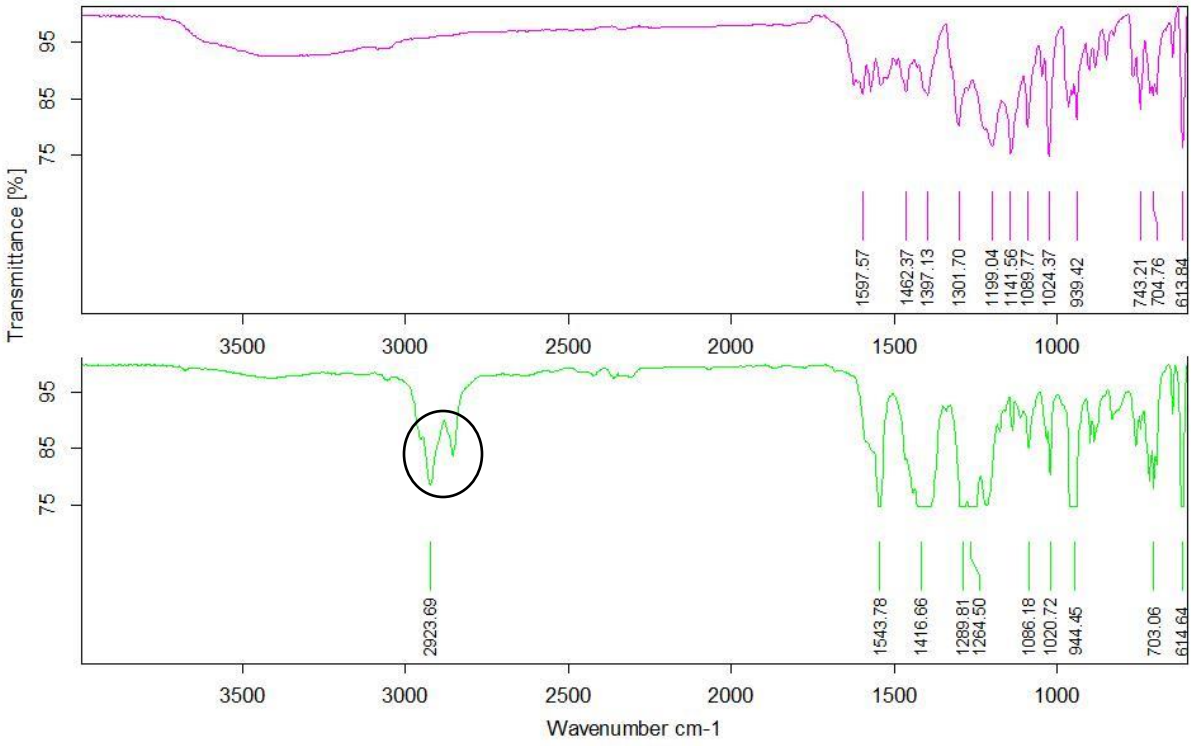


Figure S2-8A. FTIR of BPB (top) and P_2(BPB) (bottom) ; Circle- Characteristic C-H stretching confirms the formation of IL

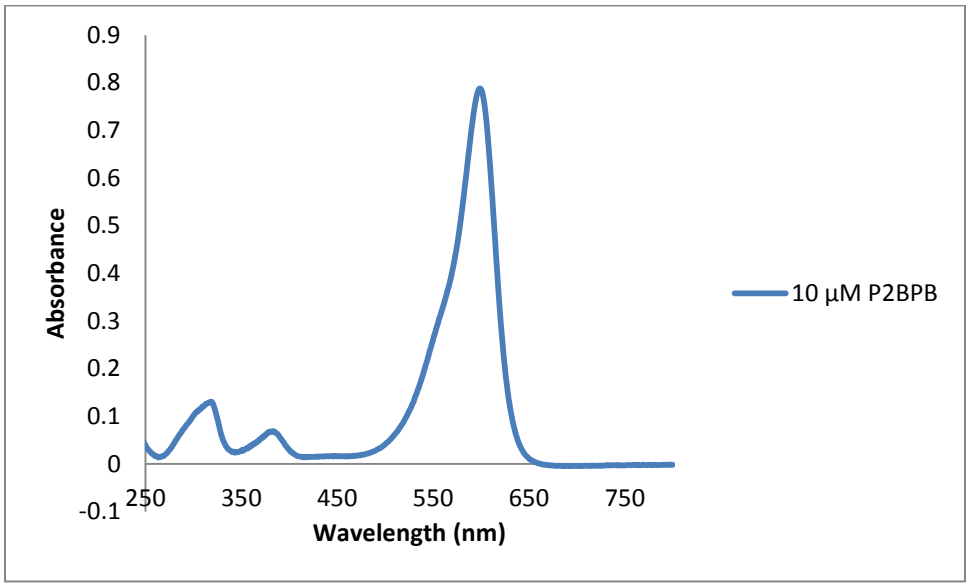


Figure S2-8B. UV-Vis spectra of P_2(BPB)

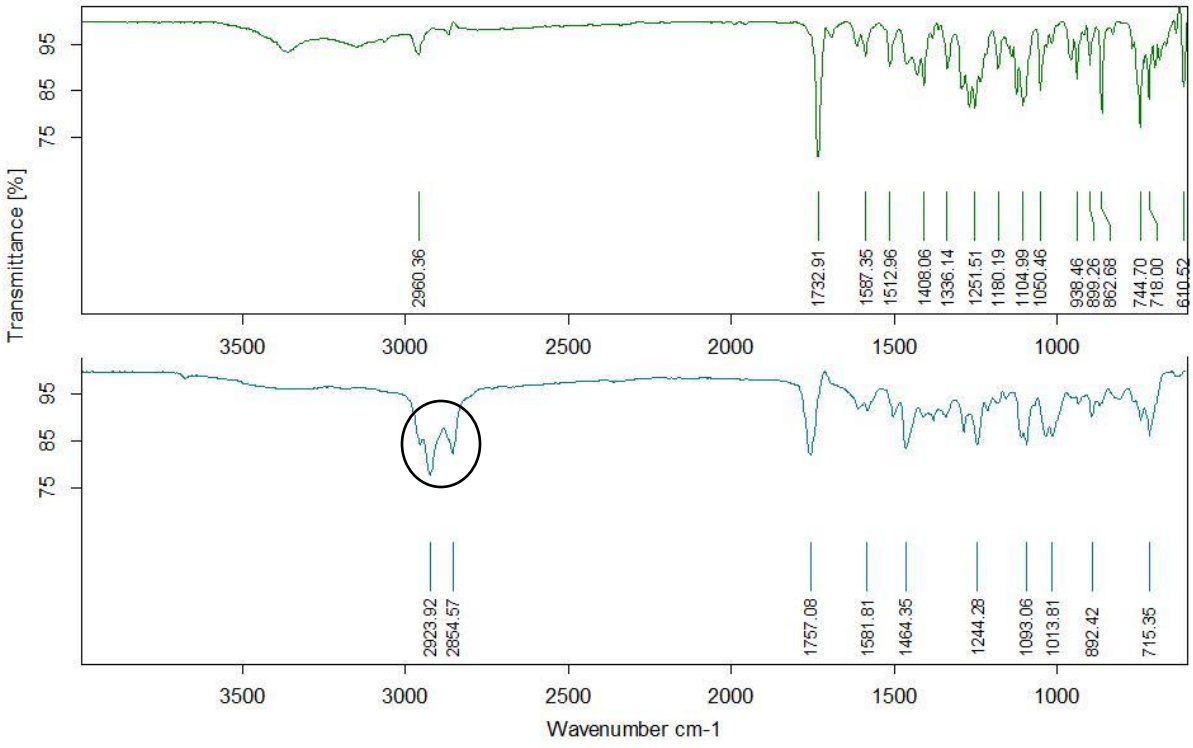


Figure S2-9A. FTIR of Thy (top) and P_2(Thy) (bottom) ; Circle- Characteristic C-H stretching confirms the formation of IL

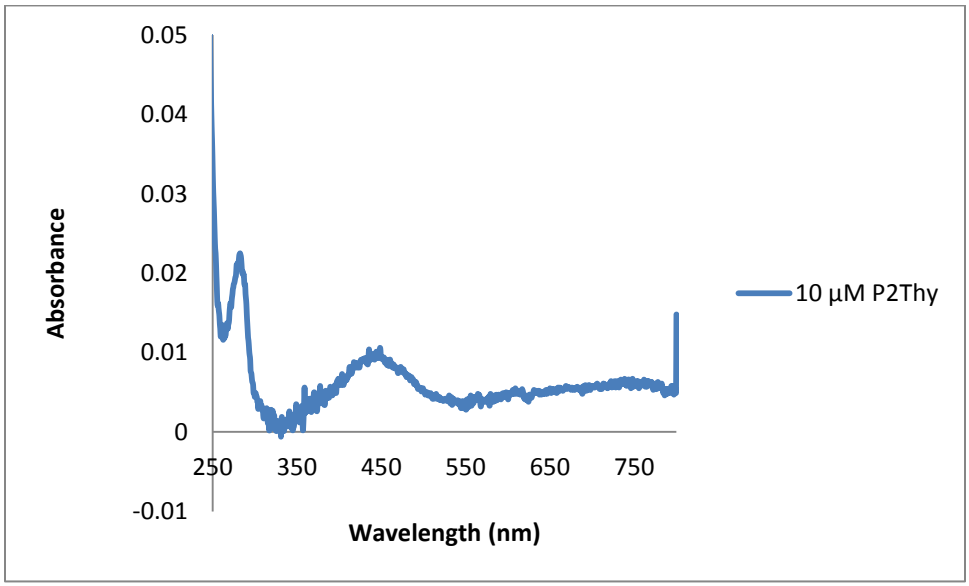


Figure S2-9B. UV-Vis spectra of P_2(Thy)

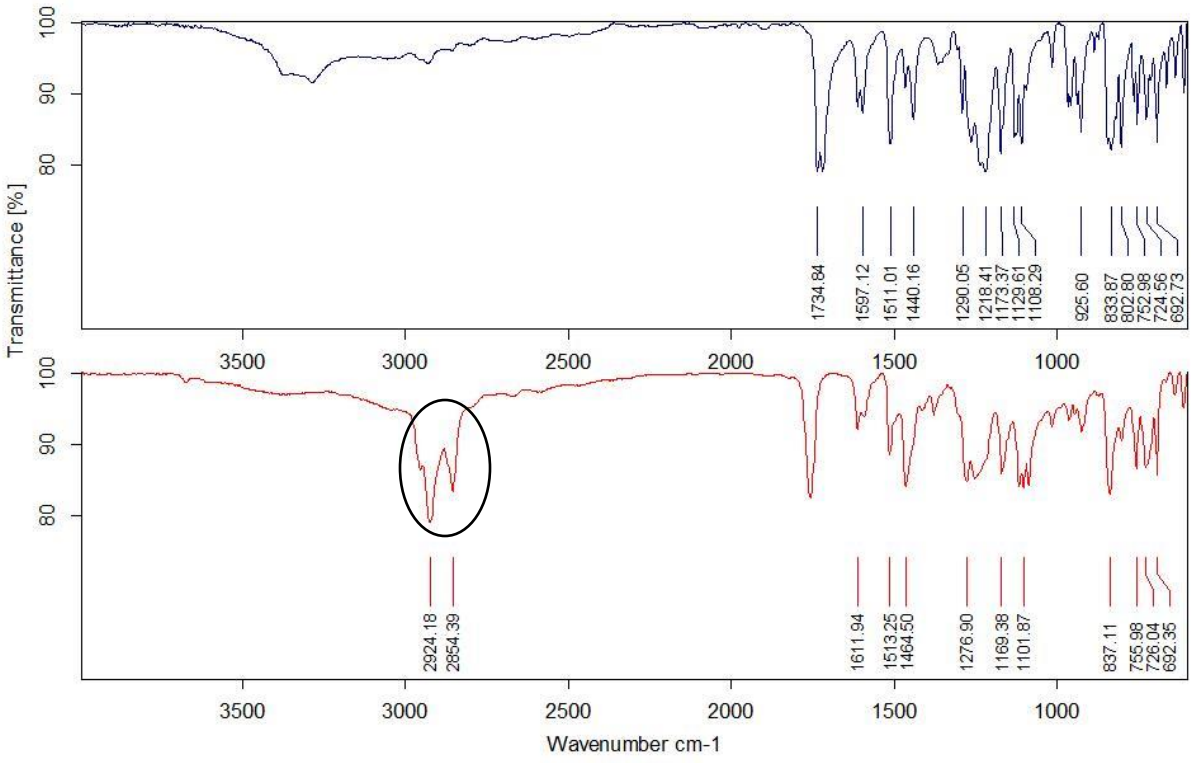


Figure S2-10A. FTIR of FFT (top) and P_2(FFT) (bottom) ; Circle- Characteristic C-H stretching confirms the formation of IL

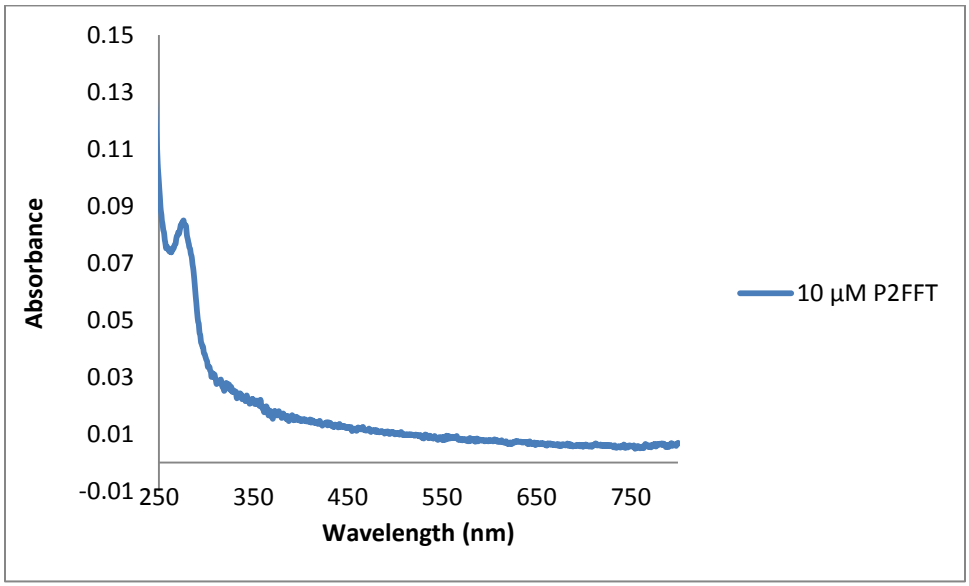


Figure S2-10B. UV-Vis spectra of P_2(FFT)

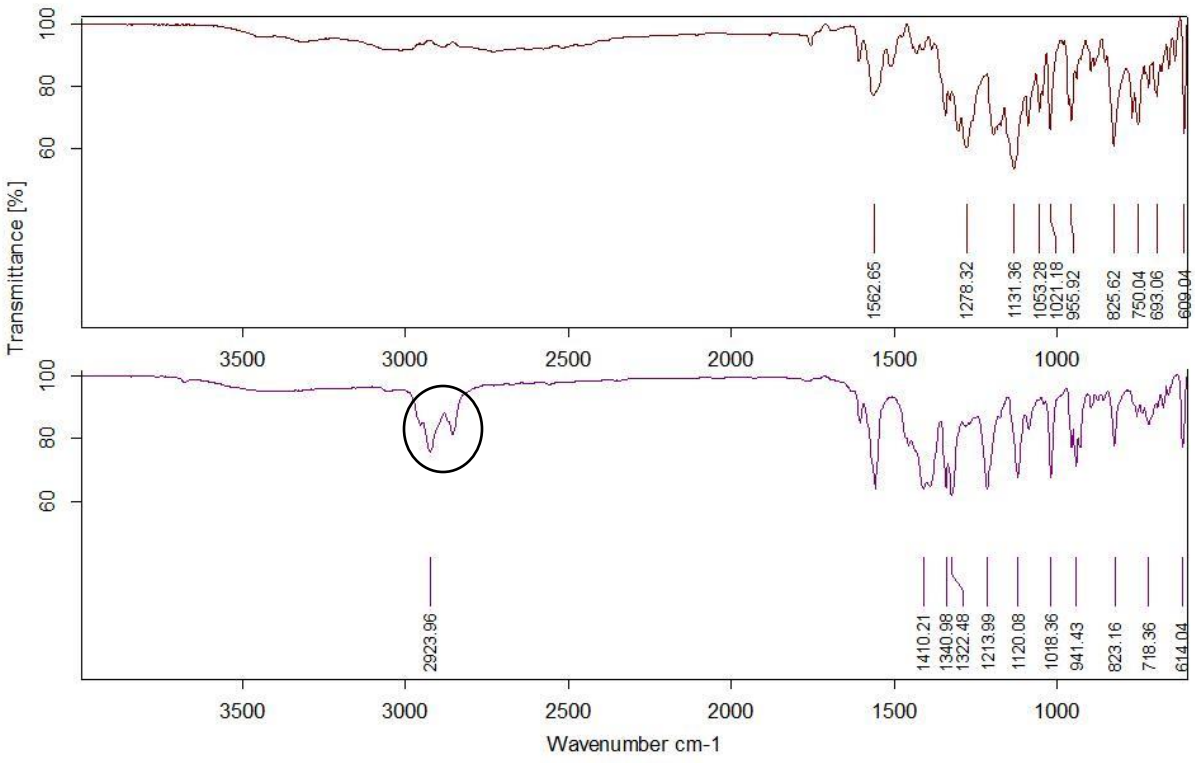


Figure S2-11A. FTIR of ClR (top) and P_2(ClR) (bottom) ; Circle- Characteristic C-H stretching confirms the formation of IL

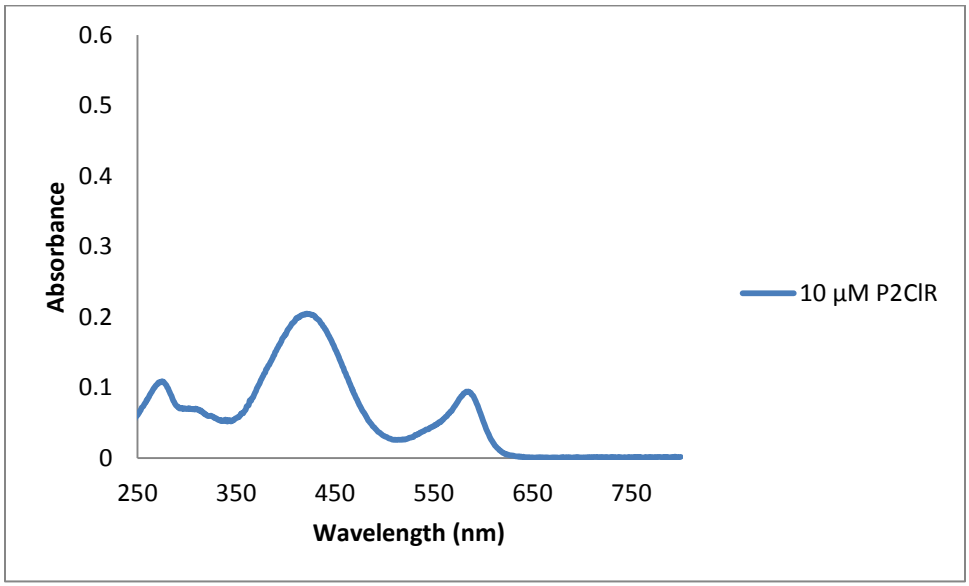


Figure S2-11B. UV-Vis spectra of P_2(ClR)

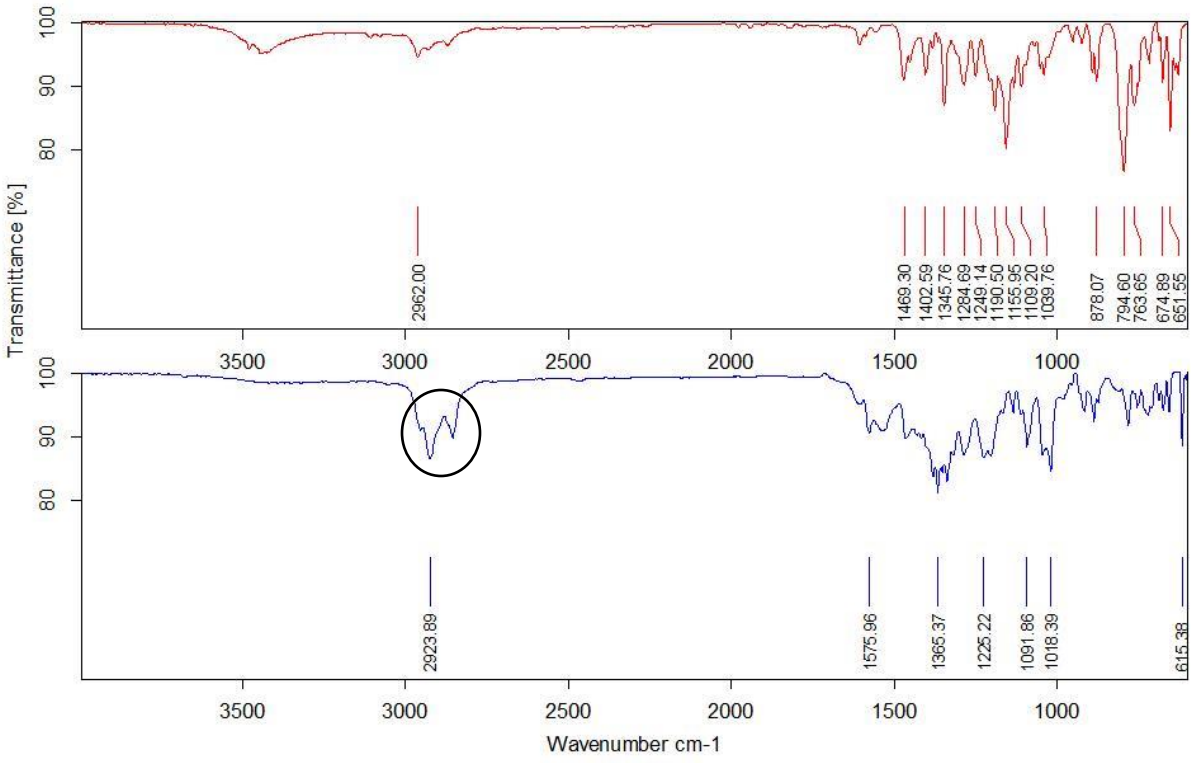


Figure S2-12A. FTIR of BTB (top) and P_2(BTB) (bottom) ; Circle- Characteristic C-H stretching confirms the formation of IL

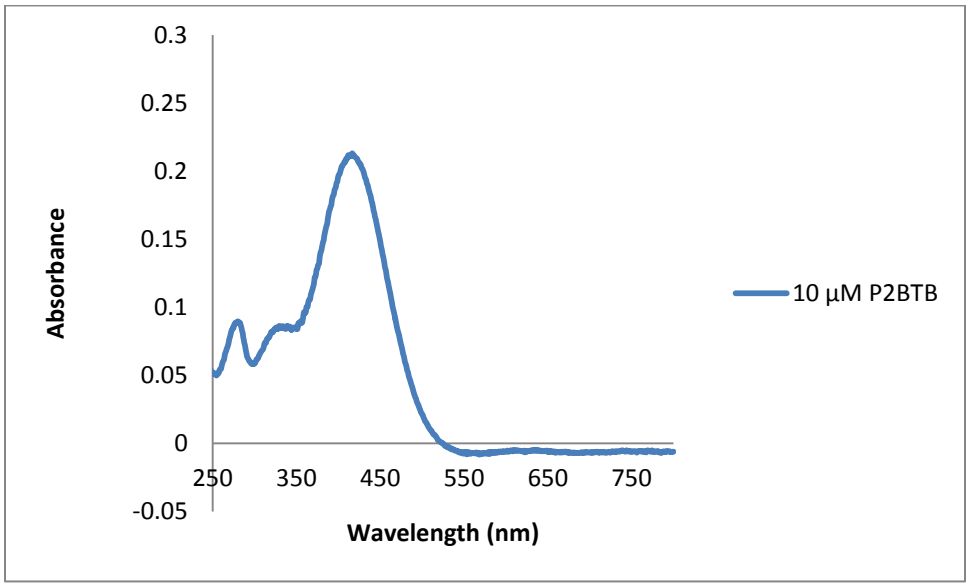


Figure S2-12B. UV-Vis spectra of P_2(BTB)

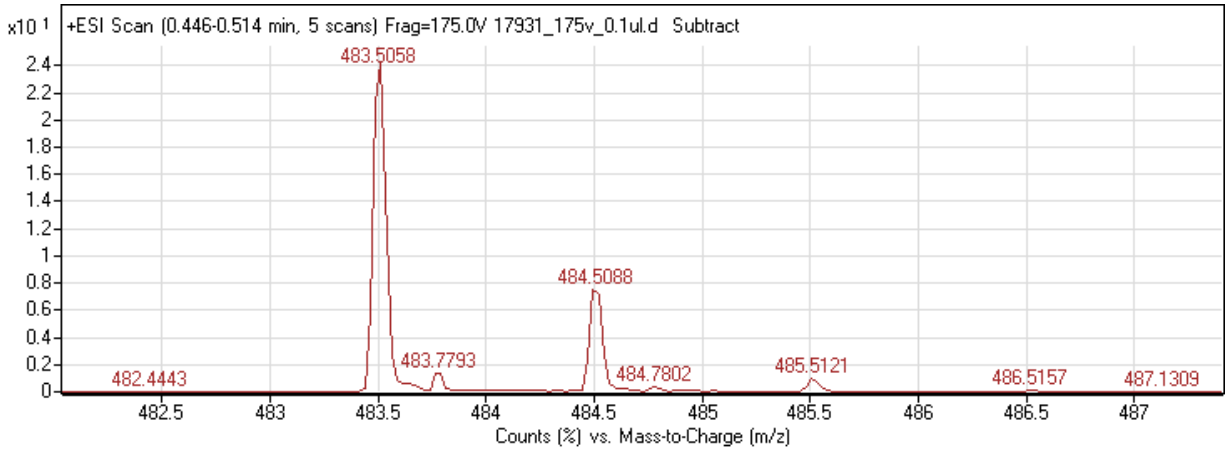


Figure S3- 1A. Electrospray ionization mass spectrum in positive ion mode P_2PR (m/z- 483.5 corresponds to $[P_{66614}]$ cation)

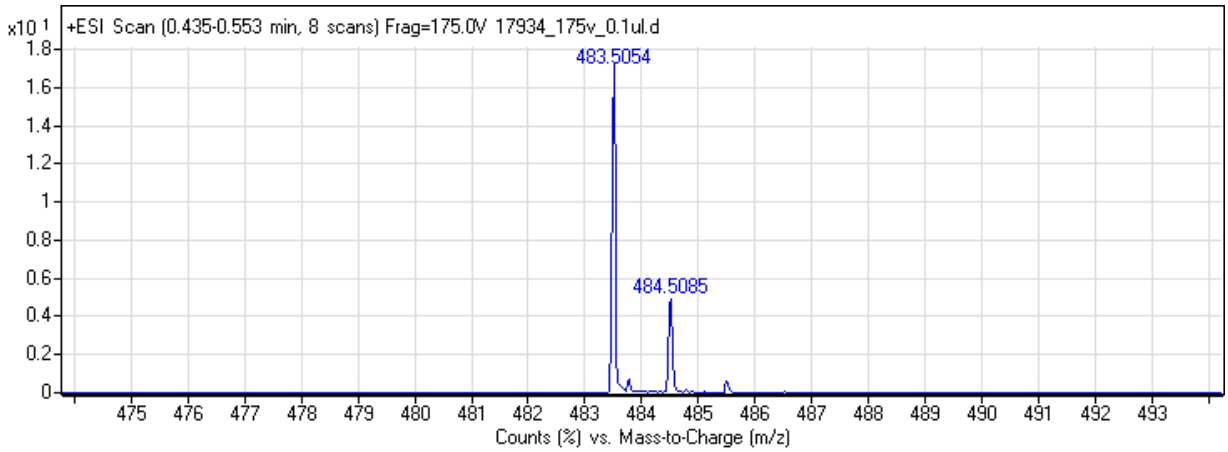


Figure S3- 2A. Electrospray ionization mass spectrum in positive ion mode P_2BY (m/z- 483.5 corresponds to $[P_{66614}]$ cation)

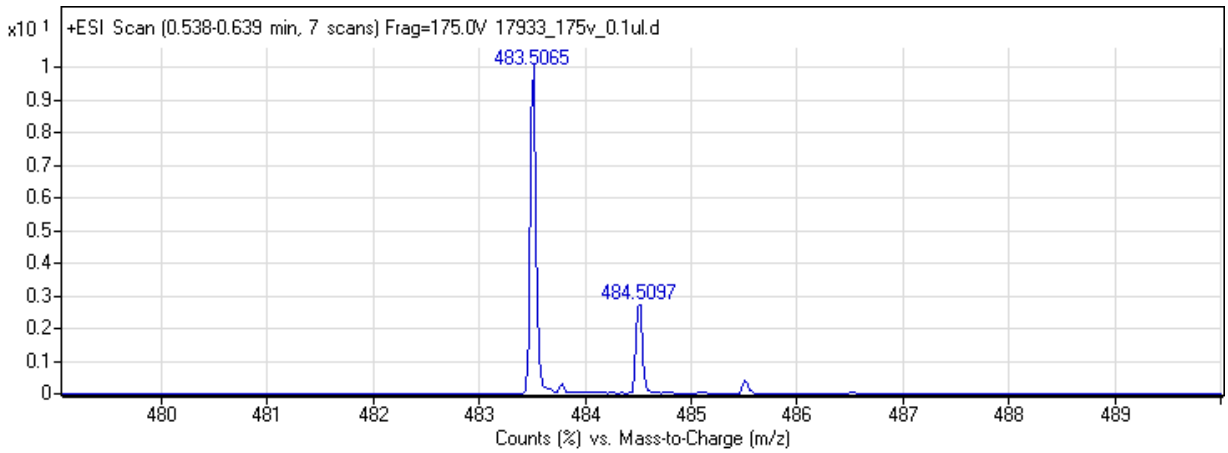


Figure S3- 3A. Electrospray ionization mass spectrum in positive ion mode P_2BCG (m/z- 483.5 corresponds to [P_{66614}] cation)

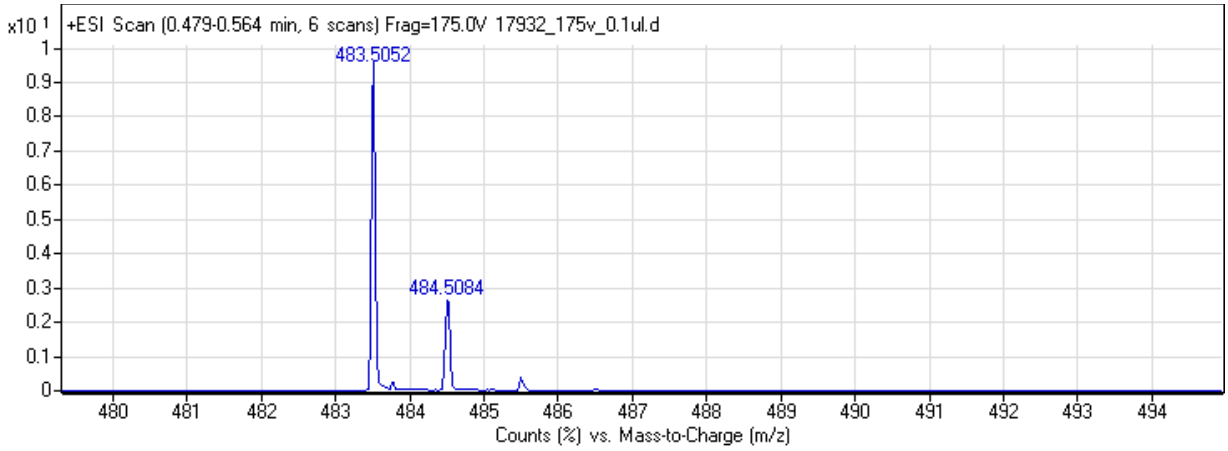


Figure S3- 4A. Electrospray ionization mass spectrum in positive ion mode P_2mCP (m/z- 483.5 corresponds to [P_{66614}] cation)

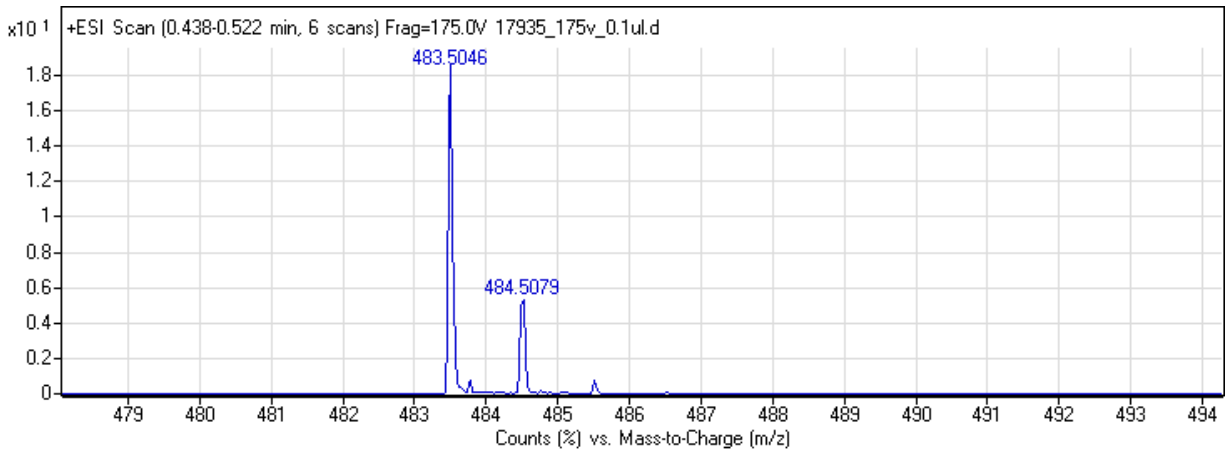


Figure S3- 5A. Electrospray ionization mass spectrum in positive ion mode PMO (m/z- 483.5 corresponds to [P_{66614}] cation)

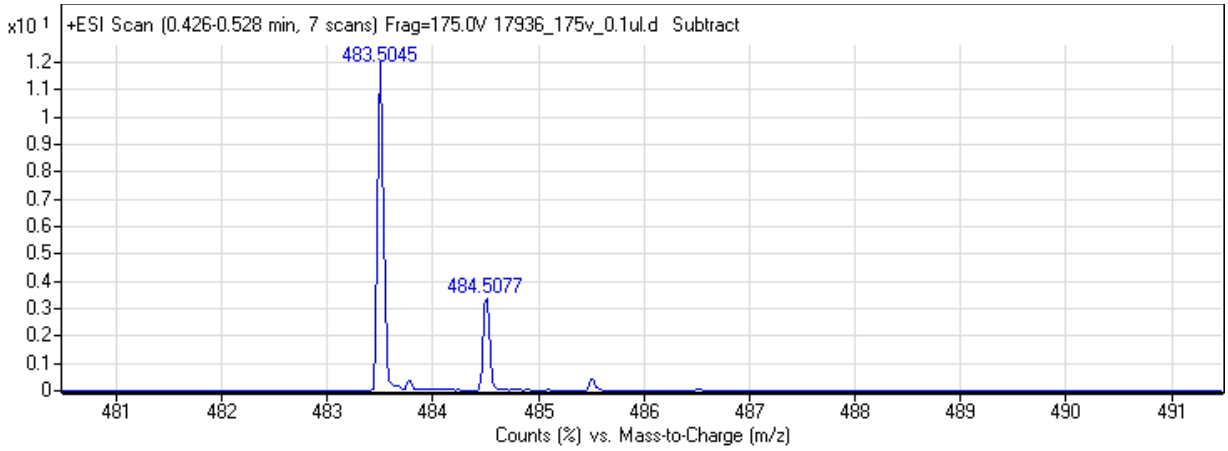


Figure S3- 6A. Electrospray ionization mass spectrum in positive ion mode PMR (m/z- 483.5 corresponds to [P_{66614}] cation)

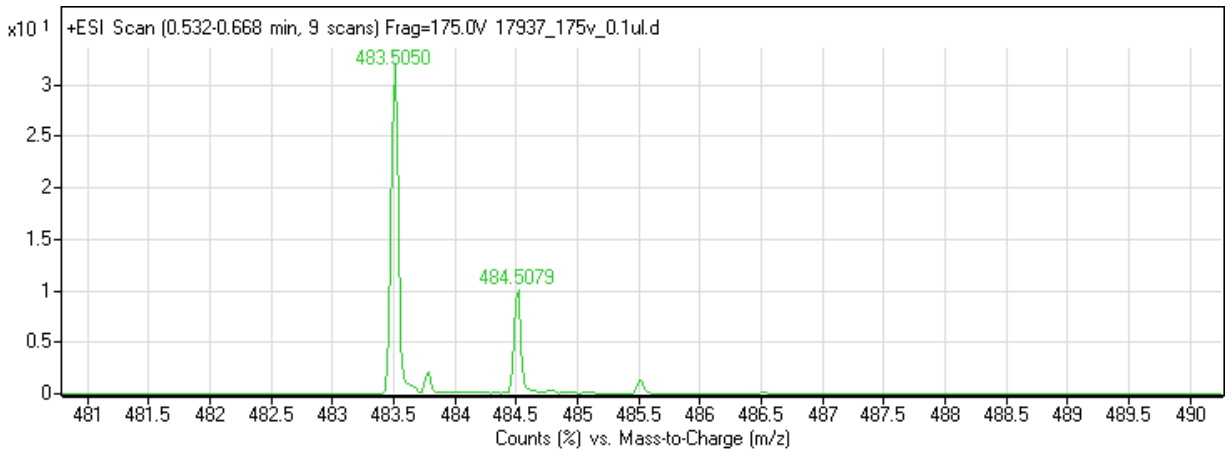


Figure S3- 7A. Electrospray ionization mass spectrum in positive ion mode P_2Xyl (m/z- 483.5 corresponds to [P_{66614}] cation)

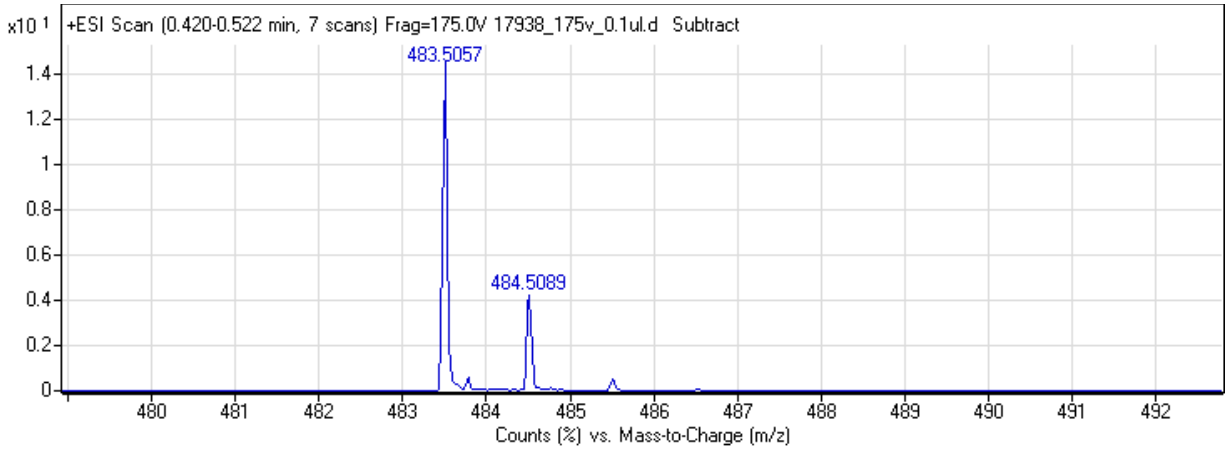


Figure S3- 8A. Electrospray ionization mass spectrum in positive ion mode P_2BPB (m/z- 483.5 corresponds to [P_{66614}] cation)

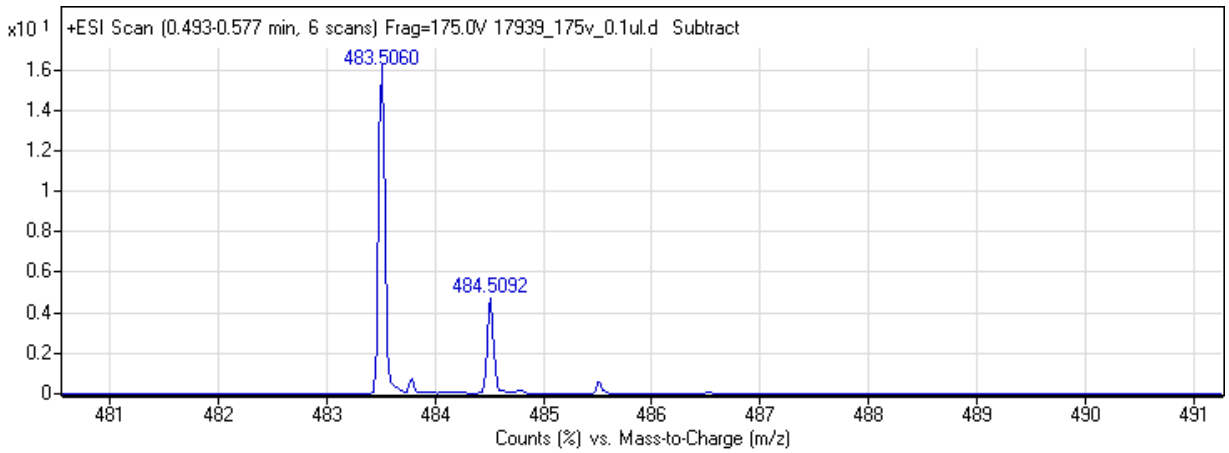


Figure S3- 9A. Electrospray ionization mass spectrum in positive ion mode P_2Thy (m/z- 483.5 corresponds to [P_{66614}] cation)

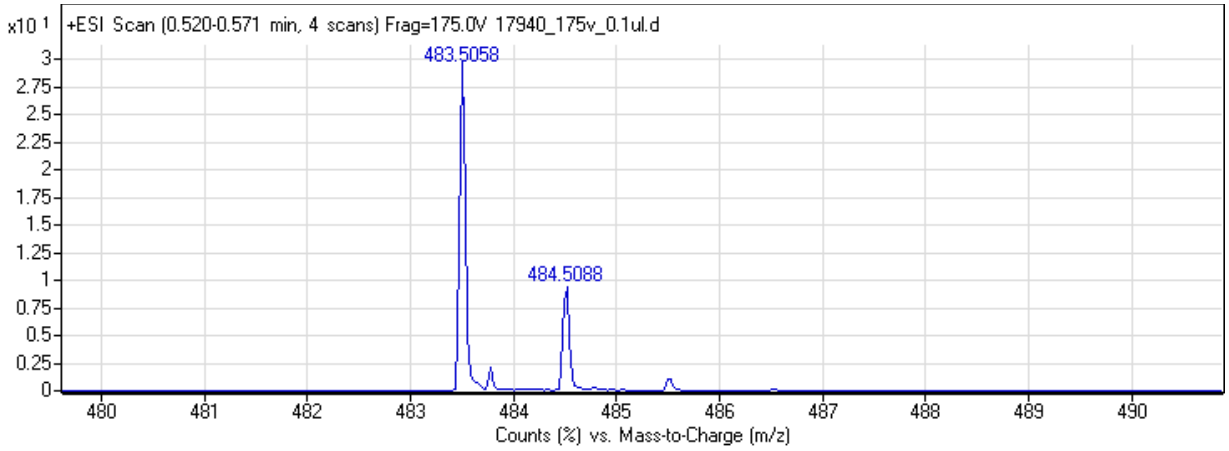


Figure S3- 10A. Electrospray ionization mass spectrum in positive ion mode P_2FFT (m/z- 483.5 corresponds to [P_{66614}] cation)

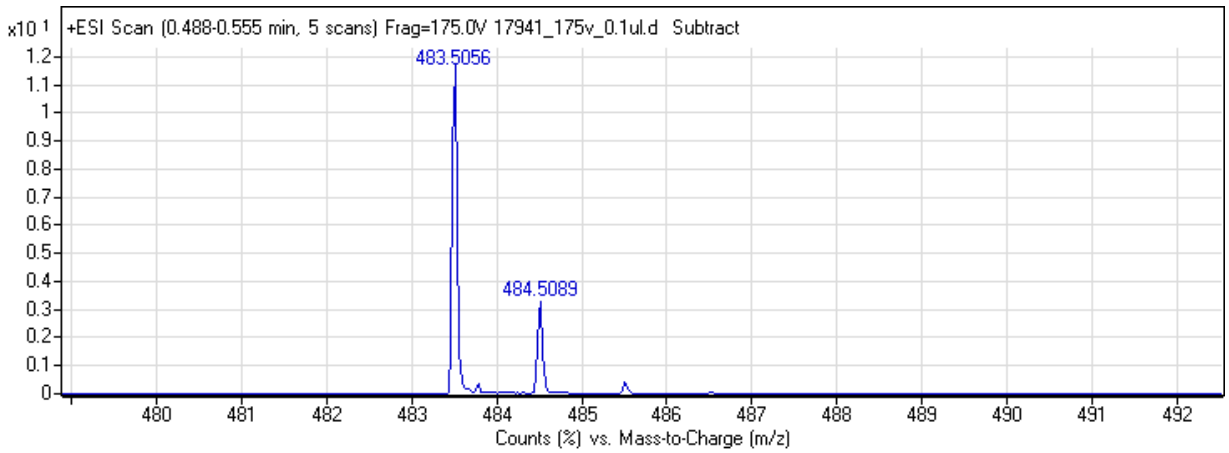


Figure S3- 11A. Electrospray ionization mass spectrum in positive ion mode P_2ClR (m/z- 483.5 corresponds to $[P_{66614}]$ cation)

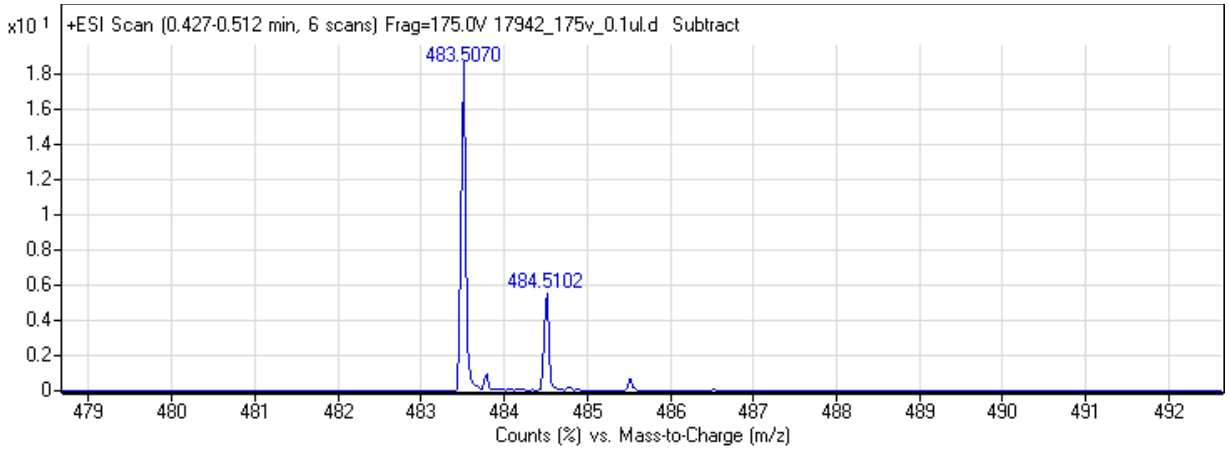


Figure S3- 12A. Electrospray ionization mass spectrum in positive ion mode P_2BTB (m/z- 483.5 corresponds to $[P_{66614}]$ cation)

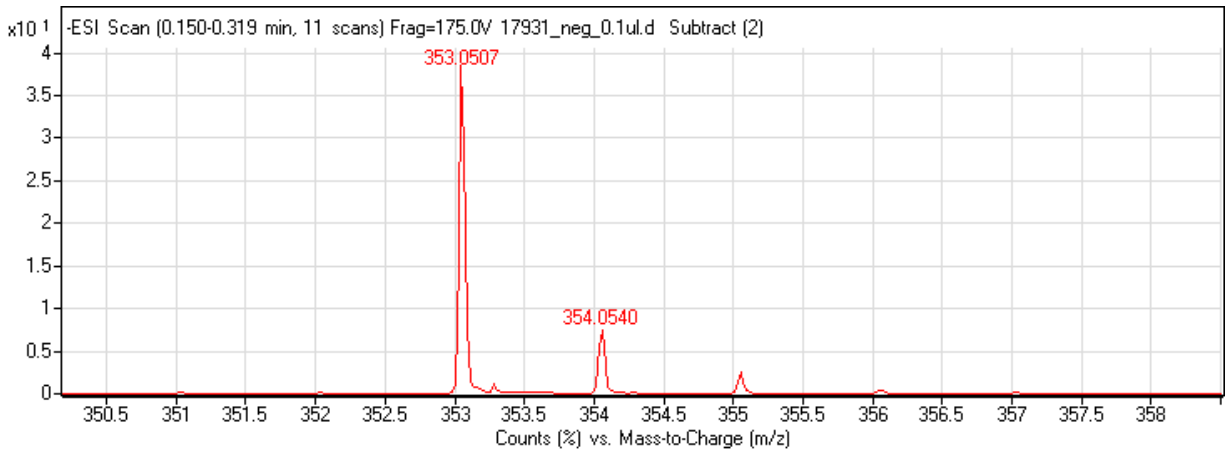


Figure S3- 1B. Electrospray ionization mass spectrum in negative ion mode P_2PR (m/z- 353.0 corresponds to PR anion)

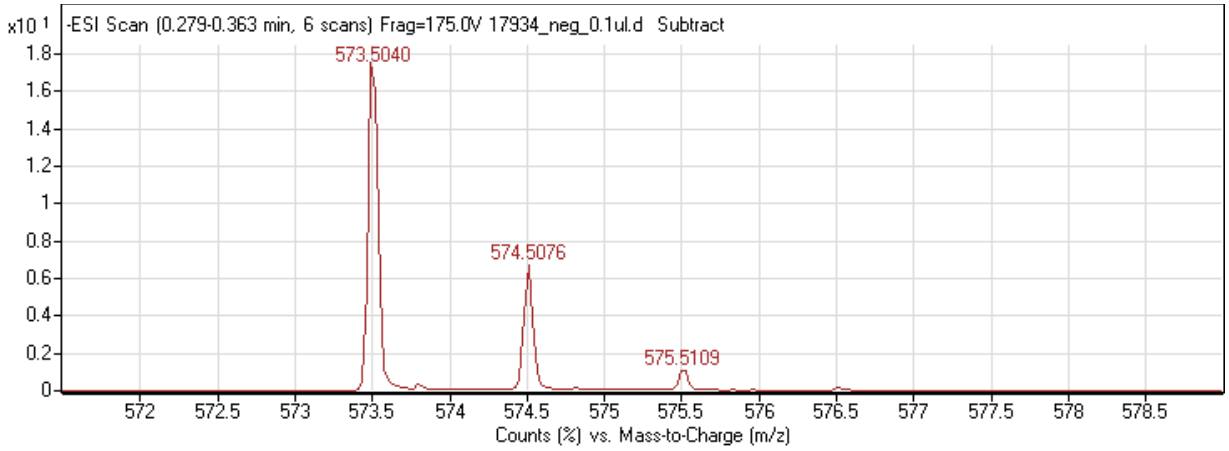


Figure S3- 2B. Electrospray ionization mass spectrum in negative ion mode P_2BY (m/z- 573.5 corresponds to BY anion)

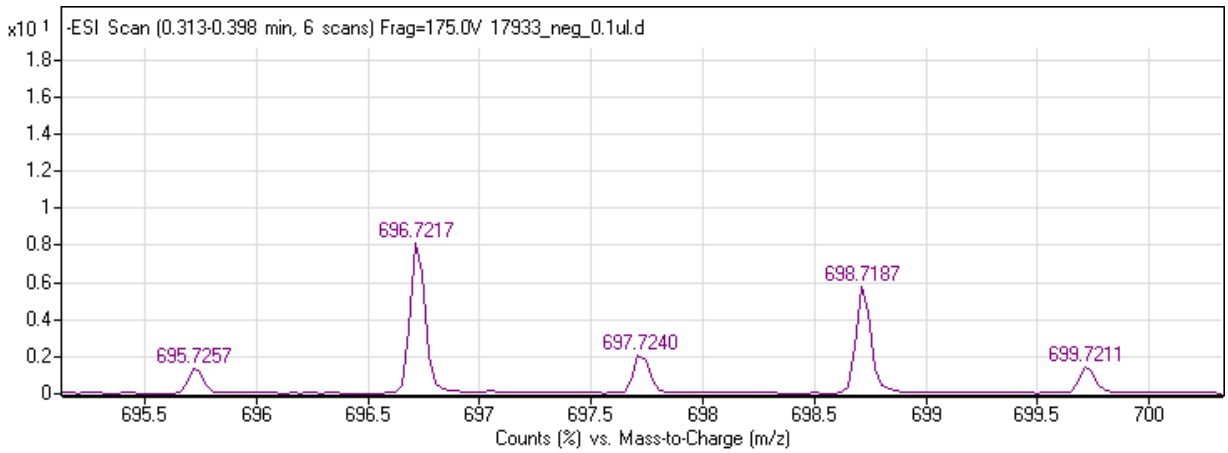


Figure S3- 3B. Electrospray ionization mass spectrum in negative ion mode P_2BCG (m/z- 696.7 corresponds to BCG anion)

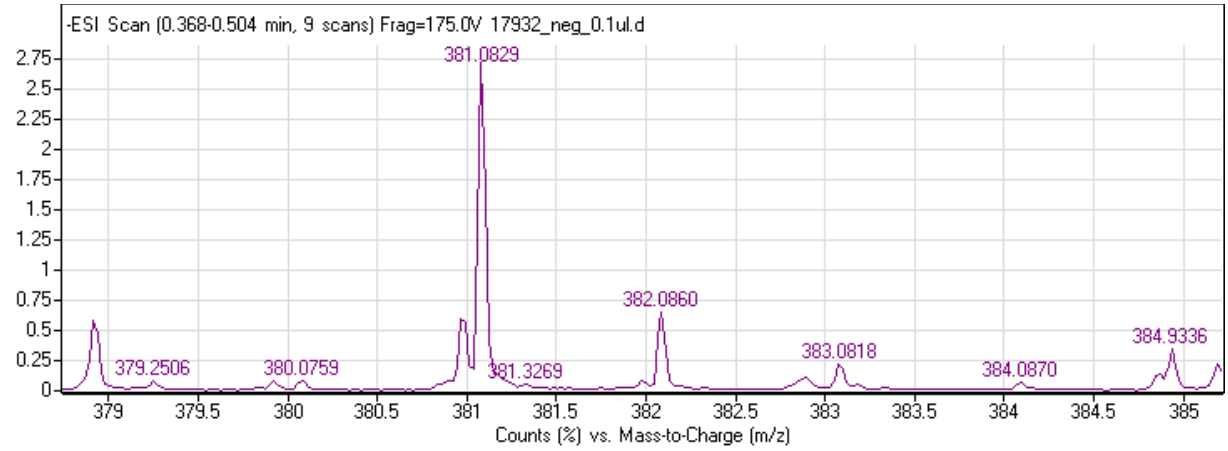


Figure S3- 4B. Electrospray ionization mass spectrum in negative ion mode P_2mCP (m/z- 381.0 corresponds to BCG anion)

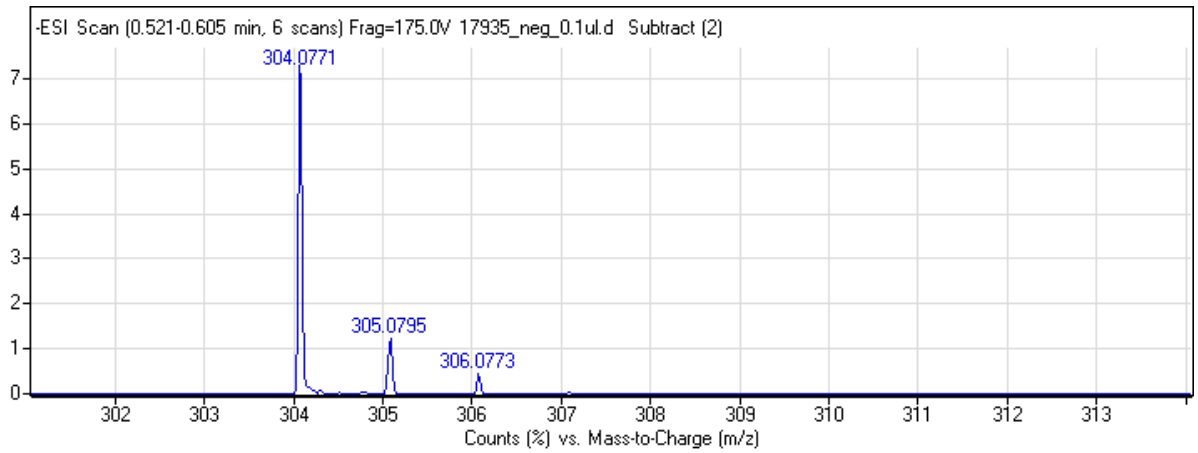


Figure S3- 5B. Electrospray ionization mass spectrum in negative ion mode PMO (m/z- 304.0 corresponds to MO anion)

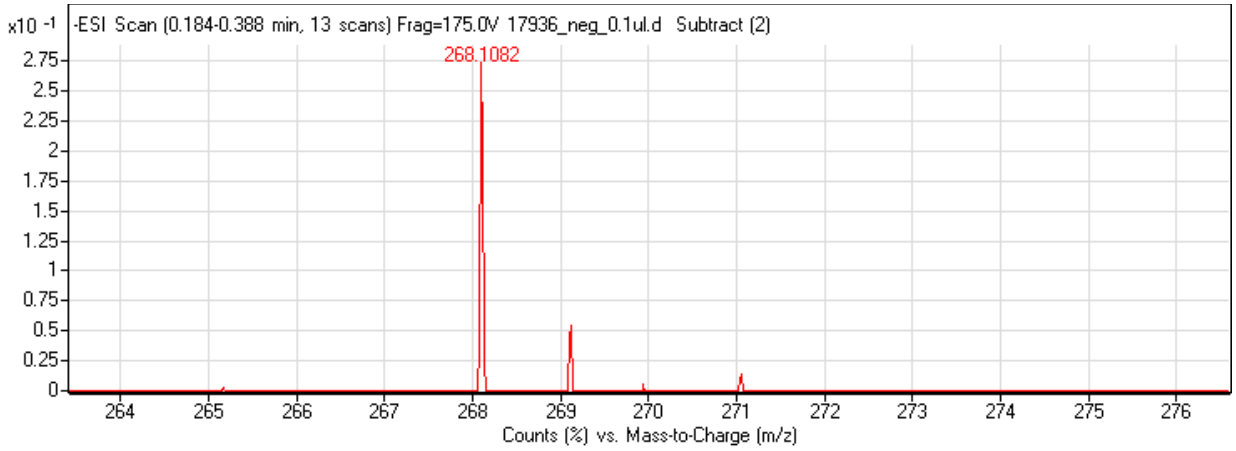


Figure S3- 6B. Electrospray ionization mass spectrum in negative ion mode PMR (m/z- 268.1 corresponds to MR anion)

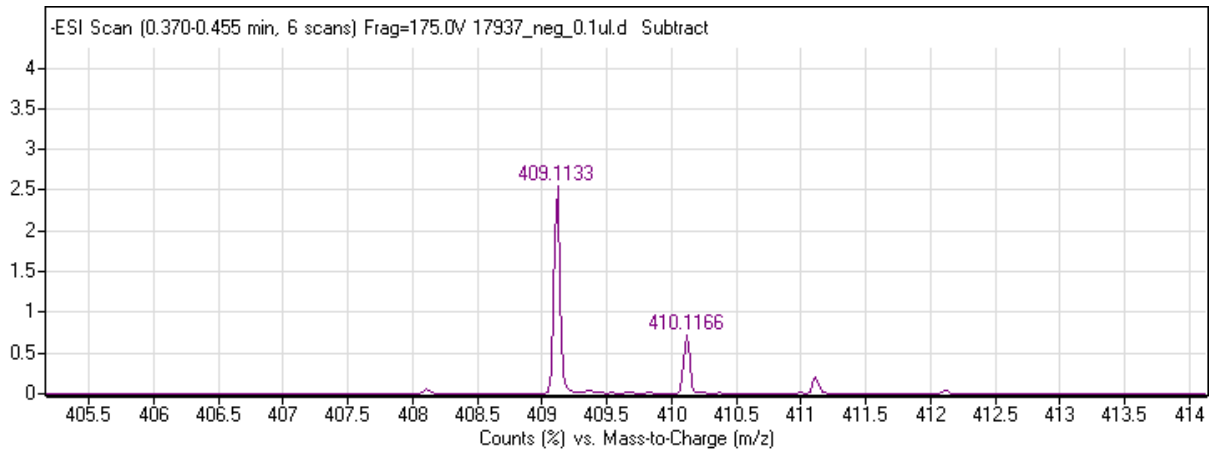


Figure S3- 7B. Electrospray ionization mass spectrum in negative ion mode P_2Xyl (m/z- 409.1 corresponds to Xyl anion)

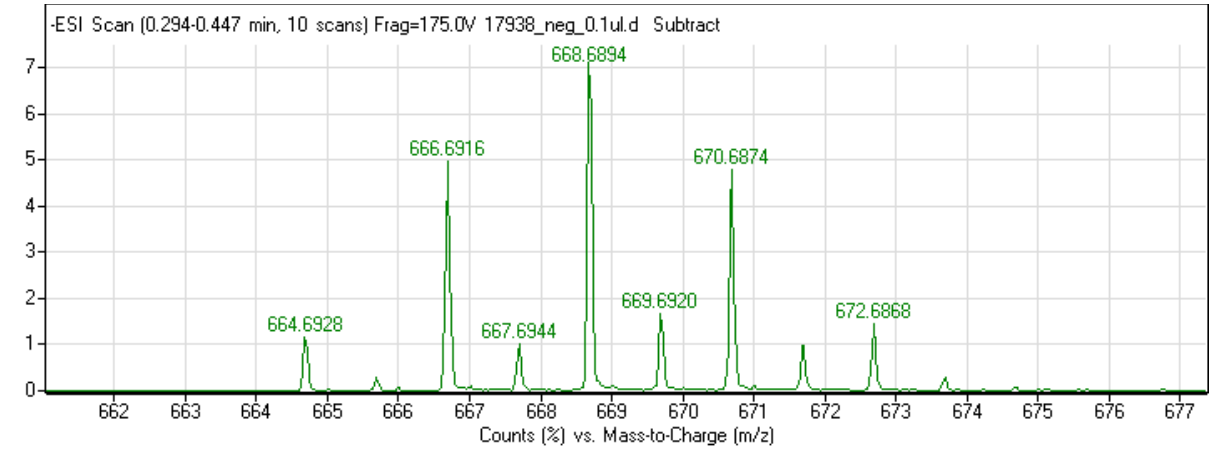


Figure S3- 8B. Electrospray ionization mass spectrum in negative ion mode P_2BPB (m/z- 668.6 corresponds to BPB anion)

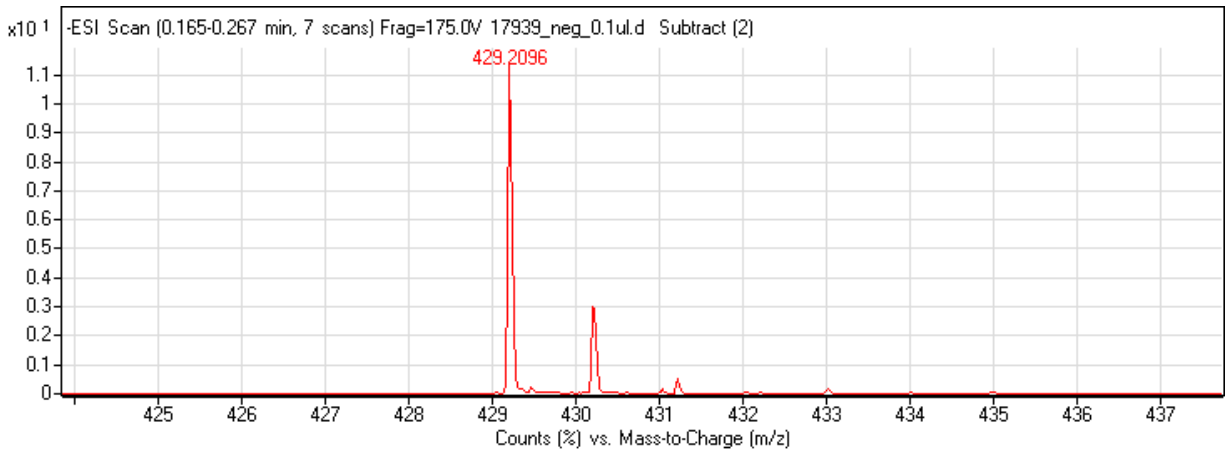


Figure S3- 9B. Electrospray ionization mass spectrum in negative ion mode P_2Thy (m/z- 429 corresponds to Thy anion)

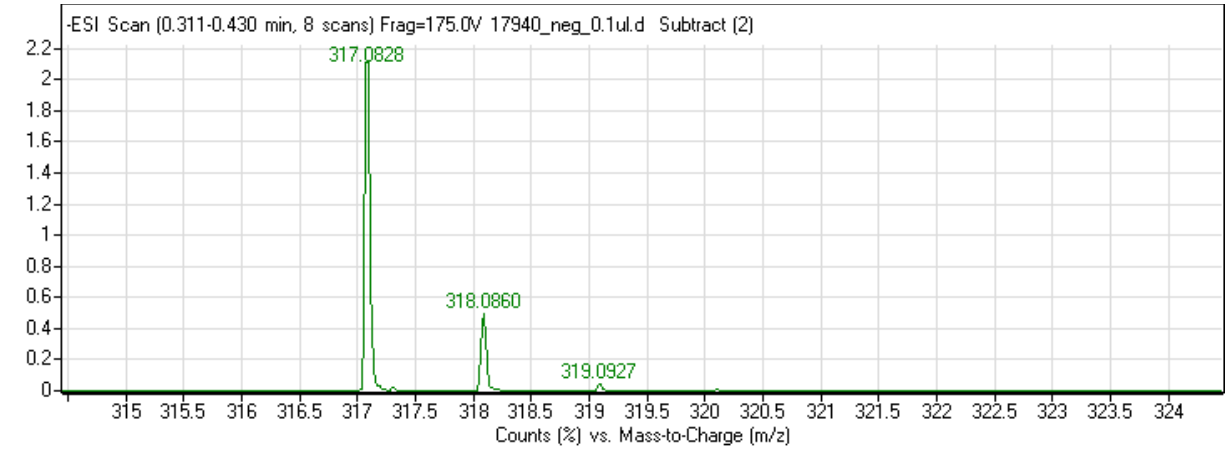


Figure S3- 10B. Electrospray ionization mass spectrum in negative ion mode P_2FFT (m/z- 317.0 corresponds to FFT anion)

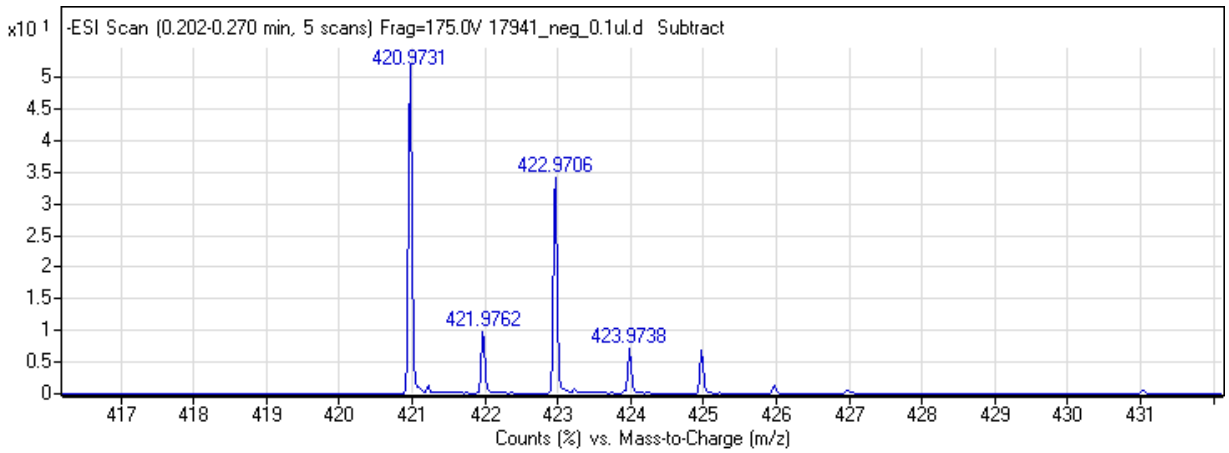


Figure S3- 11B. Electrospray ionization mass spectrum in negative ion mode P_2ClR (m/z- 420.9 corresponds to ClR anion)

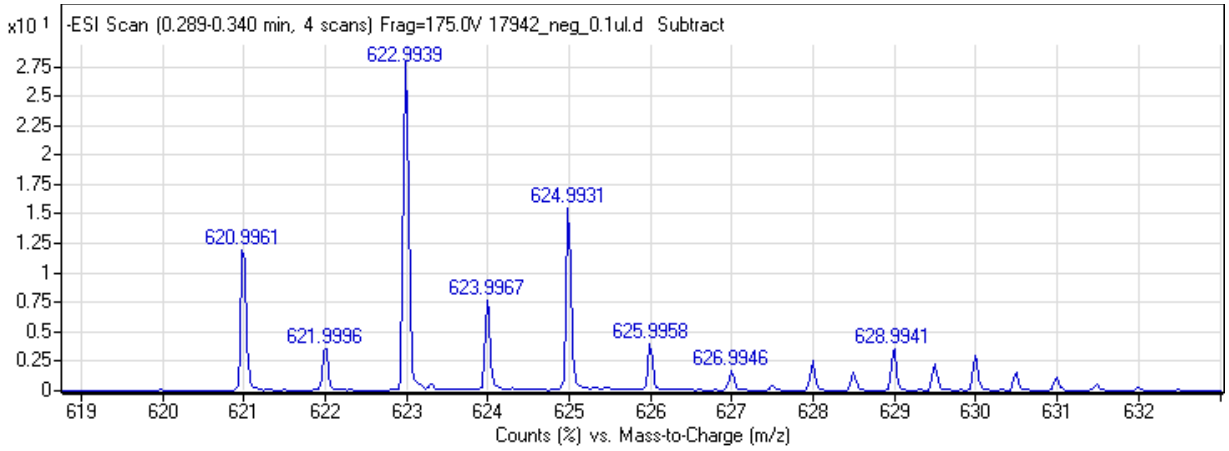


Figure S3- 12B. Electrospray ionization mass spectrum in negative ion mode P_2BTB (m/z- 622.9 corresponds to BTB anion)

Figure S4- A. Appearance of P_2(PR)

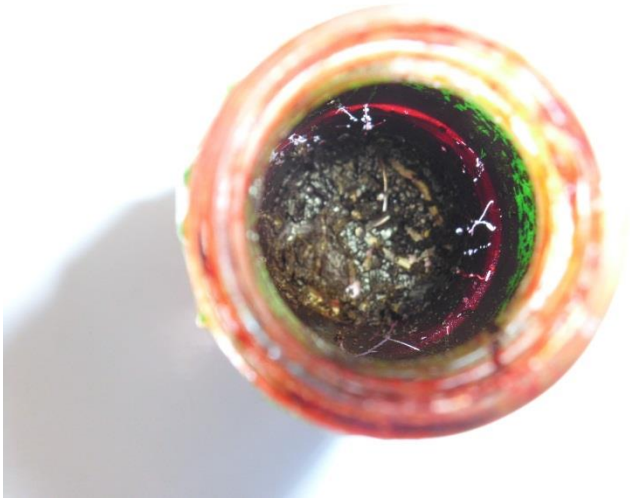


Figure S4- B. Appearance of P_2(BY)

Figure S4- C. Appearance of P_2(BCG)

Figure S4- D. Appearance of P_2(mCP)

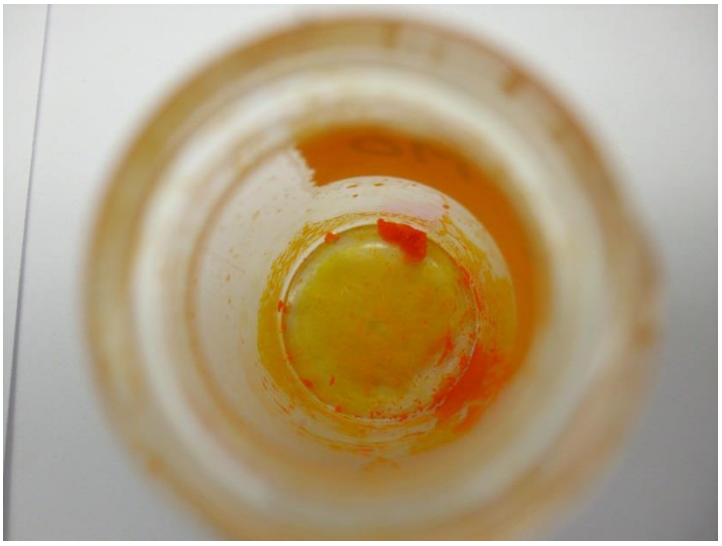


Figure S4- E. Appearance of P(MO)

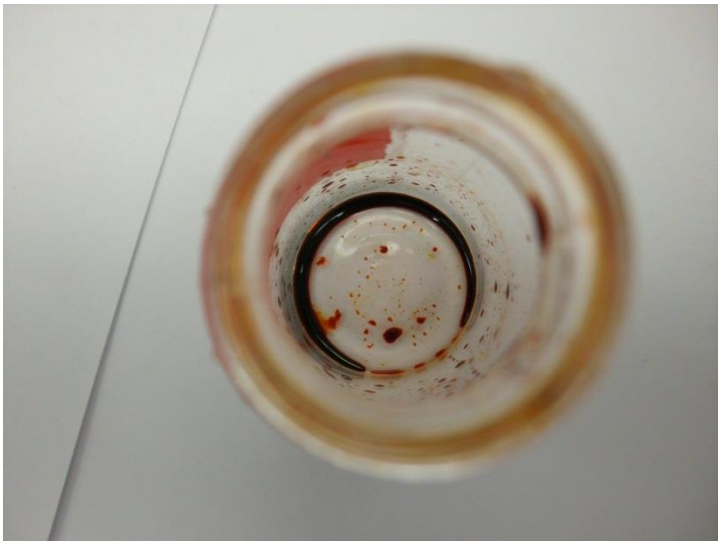


Figure S4- F. Appearance of P(MR)

Figure S4- G. Appearance of P_2(Xyl)

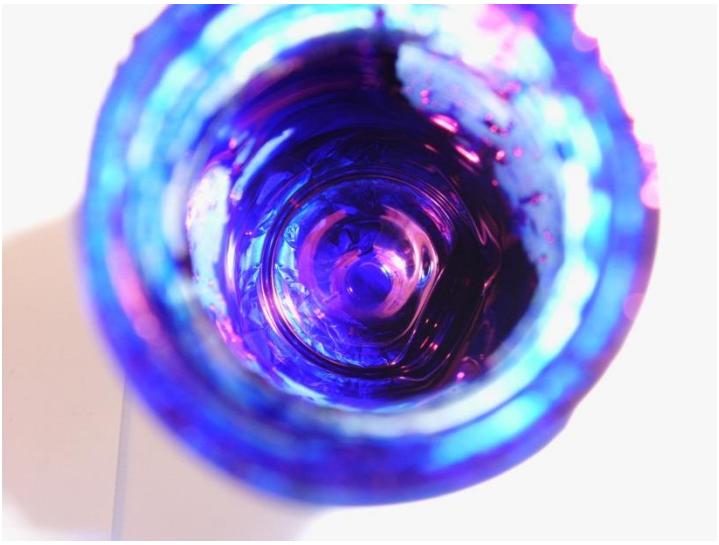


Figure S4- H. Appearance of P_2(BPB)

Figure S4- I. Appearance of P_2(Thy)

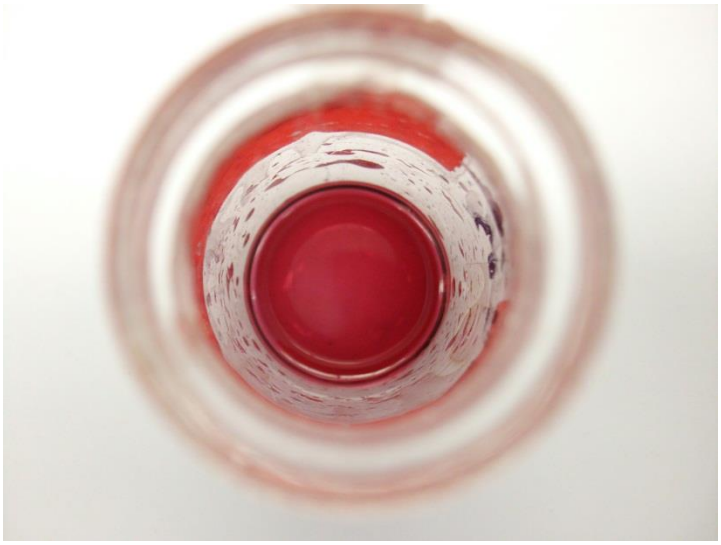
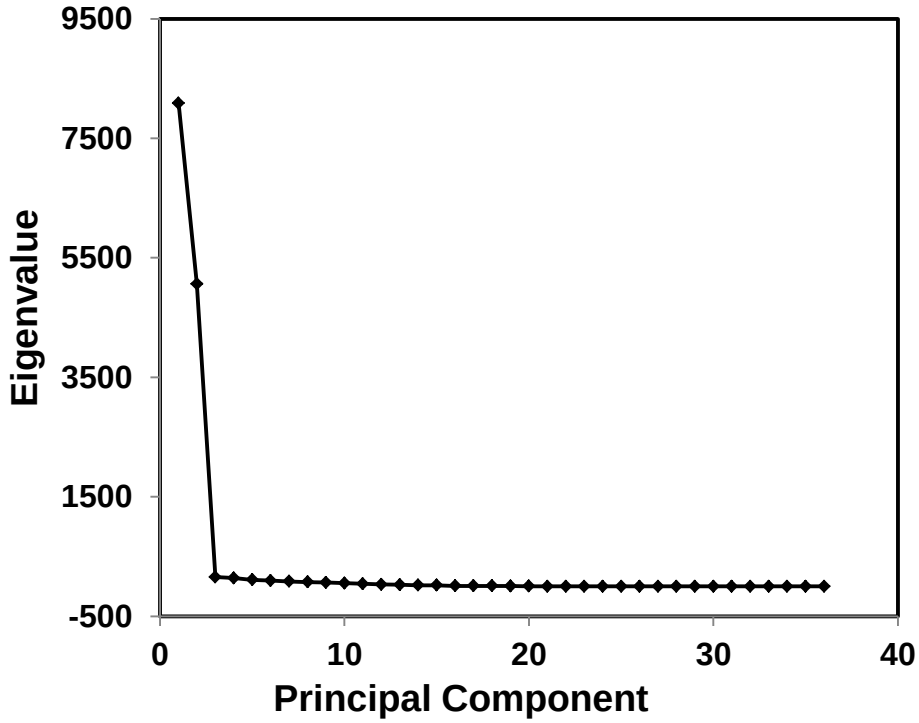


Figure S4- J. Appearance of P_2(FFT)

Figure S4- K. Appearance of P_2(ClR)

Figure S4- L. Appearance of P_2(BTB)

A



B

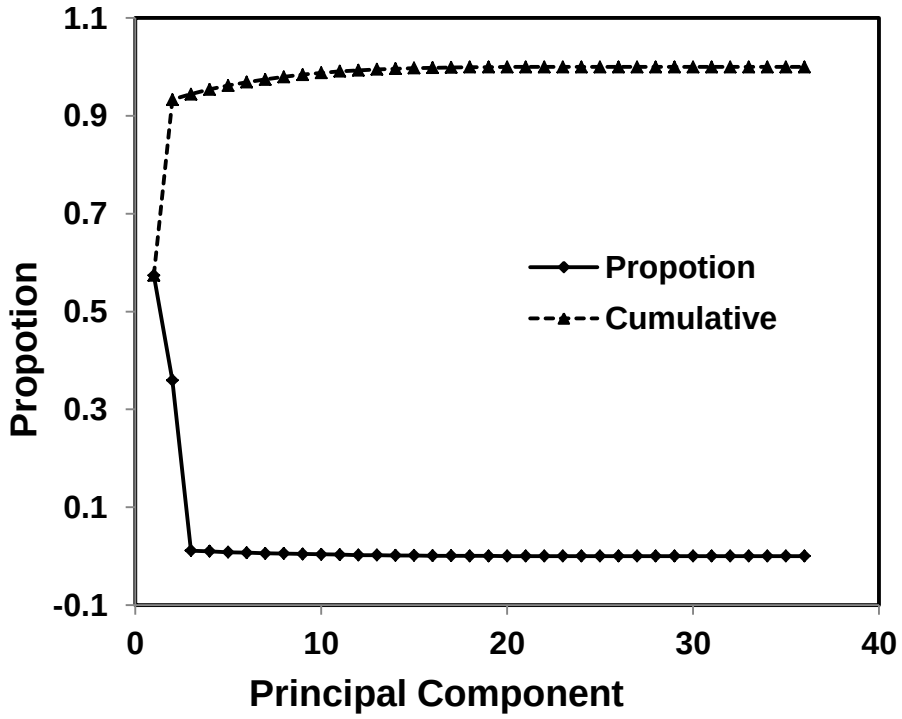


Figure S5 (A)Scree plot and (B) cumulative proportion of variability accounted for by the principal components obtained from the color change profile for the identification of cigarette smoke of Marlboro® red, Crowns® and Camel® by using filter paper based ionic liquid sensor arrays.

(1) Institute, S.: ***SAS/STAT 9.3 user's guide*****; SAS Institute, 2011.**

(2) Johnson, D. E.: ***Applied multivariate methods for data analysts*****; Duxbury Press: Pacific Grove, Calif., 1998.**