

Exploring the Relationship between Polymer Surface Chemistry and Bacterial Attachment Using ToF-SIMS and Self-Organizing maps

See Yoong Wong, Andrew L. Hook,* Wil Gardner, Chien-Yi Chang, Ying Mei, Martyn C. Davies, Paul Williams, Morgan R. Alexander, Davide Ballabio, Benjamin W. Muir, David A. Winkler, and Paul J. Pigram*

Biofilm formation is a major cause of hospital-acquired infections. Research into biofilm-resistant materials is therefore critical to reduce the frequency of these events. Polymer microarrays offer a high-throughput approach to enable the efficient discovery of novel biofilm-resistant polymers. Herein, bacterial attachment and surface chemistry are studied for a polymer microarray to improve the understanding of *Pseudomonas aeruginosa* biofilm formation on a diverse set of polymeric surfaces. The relationships between time-of-flight secondary ion mass spectrometry (ToF-SIMS) data and biofilm formation are analyzed using linear multivariate analysis (partial least squares [PLS] regression) and a nonlinear self-organizing map (SOM). The SOM models revealed several combinations of fragment ions that are positively or negatively associated with bacterial biofilm formation, which are not identified by PLS. With these insights, a second PLS model is calculated, in which interactions between key fragments (identified by the SOM) are explicitly considered. Inclusion of these terms improved the PLS model performance and shows that, without such terms, certain key fragment ions correlated with bacterial attachment may not be identified. The chemical insights provided by the combination of PLS regression and SOM will be useful for the design of materials that support negligible pathogen attachment.

1. Introduction

Nosocomial bacterial infections (hospital-acquired infections) that are resistant to standard antimicrobial treatments are the result of biofilm formation 80% of the time, as reported by Davies.^[1,2] It is the most frequent adverse event in hospital and requires urgent control measures to be implemented for implanted and indwelling medical devices. Most strategies for reducing biofilm-associated infections focus on the modification of existing materials used to manufacture indwelling medical devices via introduction of anti-biofilm compounds such as antibiotics.^[3] For example, polyurethane central-venous catheters impregnated with minocycline and rifampicin reduced bacteria colonization by 18% and prevented bloodstream infections.^[4,5] Non-antibiotic approaches with various efficacies have also been described, including treatments employing silver sulfadiazine,

S. Y. Wong, W. Gardner, P. J. Pigram
Centre for Materials and Surface Science and the Department of
Mathematical and Physical Sciences
La Trobe University
Melbourne, Victoria 3086, Australia
E-mail: p.pigram@latrobe.edu.au

A. L. Hook, M. C. Davies, M. R. Alexander, D. A. Winkler
Advanced Materials and Healthcare Technologies
School of Pharmacy
University of Nottingham
Nottingham NG7 2RD, UK
E-mail: andrew.hook@nottingham.ac.uk



The ORCID identification number(s) for the author(s) of this article can be found under <https://doi.org/10.1002/admi.202202334>.

© 2023 The Authors. Advanced Materials Interfaces published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

DOI: 10.1002/admi.202202334

C.-Y. Chang
School of Dental Science
Newcastle University
Newcastle upon Tyne NE2 4BW, UK

Y. Mei
Department of Bioengineering
Clemson University
Charleston, SC 29425, USA

P. Williams
Biodiscovery Institute
University of Nottingham
Nottingham NG7 2RD, UK

D. Ballabio
Milano Chemometrics and QSAR Research Group
Department of Earth and Environmental Sciences
University of Milano-Bicocca
Piazza della Scienza 1, 20126 Milano, Italy

B. W. Muir
CSIRO Manufacturing
Clayton, Victoria 3168, Australia

nitrofurazone, chlorhexidine and polymerized quaternary ammonium surfactants,^[6–11] whilst a number of strategies to disrupt biofilm have also been explored.^[12] Sustaining the efficacy of loaded devices is challenging and has limited their clinical impact.^[13–16] Materials that are inherently resistant to bacterial attachment and biofilm formation provide an alternative approach to preventing device-associated infections that avoids the development of antimicrobial resistance.^[17] However, the necessary understanding of biofilm formation on polymeric materials is not sufficiently developed to enable ab initio design of biofilm-resistant materials.

High-throughput experimental polymer microarray screening has enabled the discovery of anti-biofilm materials without the requirement of understanding the biological–material interaction. Beyond materials discovery, the hundreds of experimentally assessed biofilm formation data points acquired on these polymeric materials provide an ideal dataset to develop the understanding of how polymers influence bacterial behavior.

Time-of-flight secondary ion mass spectrometry (ToF-SIMS) provides a comprehensive assessment of the surface chemistry of polymers,^[18–20] delivering high surface sensitivity and detailed molecular insights,^[21,22] that can be used for the construction of structure–function relationships.^[23] Multivariate analysis (MVA) techniques, such as partial least squares (PLS) regression and principal component analysis (PCA) have traditionally been used to interpret large, high dimensional ToF-SIMS spectral datasets.^[24,25] MVA is useful for datasets containing linear correlations;^[26] however, many materials' structure–property relationships are nonlinear. The presence of outliers and missing data also affect the quality of MVA outcomes.^[27]

Hook et al. reported a PLS regression study of the role of weak amphiphilic properties on resistance to biofilm formation.^[28] Multiple chemical moieties were implicated in bacterial adhesion, specifically, ethylene glycol and hydroxyl side groups were found to promote bacterial adhesion whilst the combination of esters with hydrophobic side groups was shown to resist bacterial adhesion.^[28,29] The same datasets were also used to construct non-linear quantitative structure–property relationship machine learning models using a Bayesian regularized artificial neural network (ANN).^[30–32] These studies showed that nonlinear machine learning models are often more robust and predictive than conventional linear regression models. This work identified the relationships between specific features associated with hydrophobicity or proton transfer, and molecular shape and their propensity to prevent bacterial attachment to surfaces. The effect of combinations of chemical features, however, could not be completely resolved, suggesting that further study of important and complex features present in the system is needed.

D. A. Winkler
Department of Biochemistry and Chemistry
La Trobe Institute for Molecular Sciences
La Trobe University
Melbourne, Victoria 3086, Australia

D. A. Winkler
Monash Institute of Pharmaceutical Sciences
Monash University
Parkville 3052, Australia

McCulloch and Pitts first introduced the concept of threshold logic, a computational model created to mimic the biological neural system based on mathematics and algorithms, 80 years ago.^[33–35] This model was the precursor of the widely used ANNs algorithms now very much in vogue. Kohonen introduced a non-linear algorithm known as the self-organizing map (SOM), or Kohonen network, based on ANNs.^[36] Variants of the SOM have been developed, such as the counter-propagation artificial neural network (CPN), supervised Kohonen network (SKN), and the X-Y-fused Kohonen network (XYF) that are capable of predicting characteristics of interest.^[27] They have found broad application in diverse fields such as engineering,^[37] environmental science,^[38] hyperspectral imaging,^[39] and finance.^[40] The ability of SOMs to provide a visual interpretation of high dimensional data make them highly valuable in many contexts. We have previously studied the application of unsupervised SOM to acrylate polymer microarrays.^[41] SOMs showed exceptional performance in distinguishing chemically very similar polymers based on the topological relatedness of their surface chemistries.

In this study a polymer microarray has been prepared to include material properties and surface chemistries previously implicated in preventing bacterial biofilm formation, to enable an investigation of bacterial attachment to polymers using ToF-SIMS to characterize surface chemistry (**Figure 1**). The study has a particular focus on identifying where multiple factors play a role in complex, complementary, and competitive ways. A PLS regression analysis is first presented to compare the conventional analysis approach with SOMs directly. We then employ SOMs to identify relationships between polymer surface chemistry, resolved at the level of discrete molecular structures by ToF-SIMS, and functionality, namely the attachment of *Pseudomonas aeruginosa*, a pathogen frequently implicated in hospital-acquired bacterial infections.

2. Principles of the SOM

Figure 2 shows a schematic description of the SOMs employed in this study, highlighting the information flow in each type of SOM regarding sample class assignment. Using the dataset shown in **Figure 1** as an example, each weight layer (W_x) represents each secondary ion fragment selected in the ToF-SIMS spectra (**Figure 1c**), whereas each class layer (C_l) represents the degree of *P. aeruginosa* attachment acquired from the bacterial attachment assay (**Figure 1d**).

All SOMs are initialized by assigning a random number to the weight of each map layer of all network units, which are known as neurons (N_x). In other words, each neuron contains a full list secondary ion fragment selected, and a random number is assigned to each weight layer of the neuron. The UKN, illustrated as the grey block in **Figure 2**, is an unsupervised SOM where the input objects of sample *S* were presented to the entire set of network units. The neuron with weight vectors most like the sample is then labelled the winning neuron. The weight vectors of the winning neuron and its nearest neighbors are then updated according to the sample assigned to the winning neuron. This matching and adjustment process would be repeated based on the number of epochs defined by

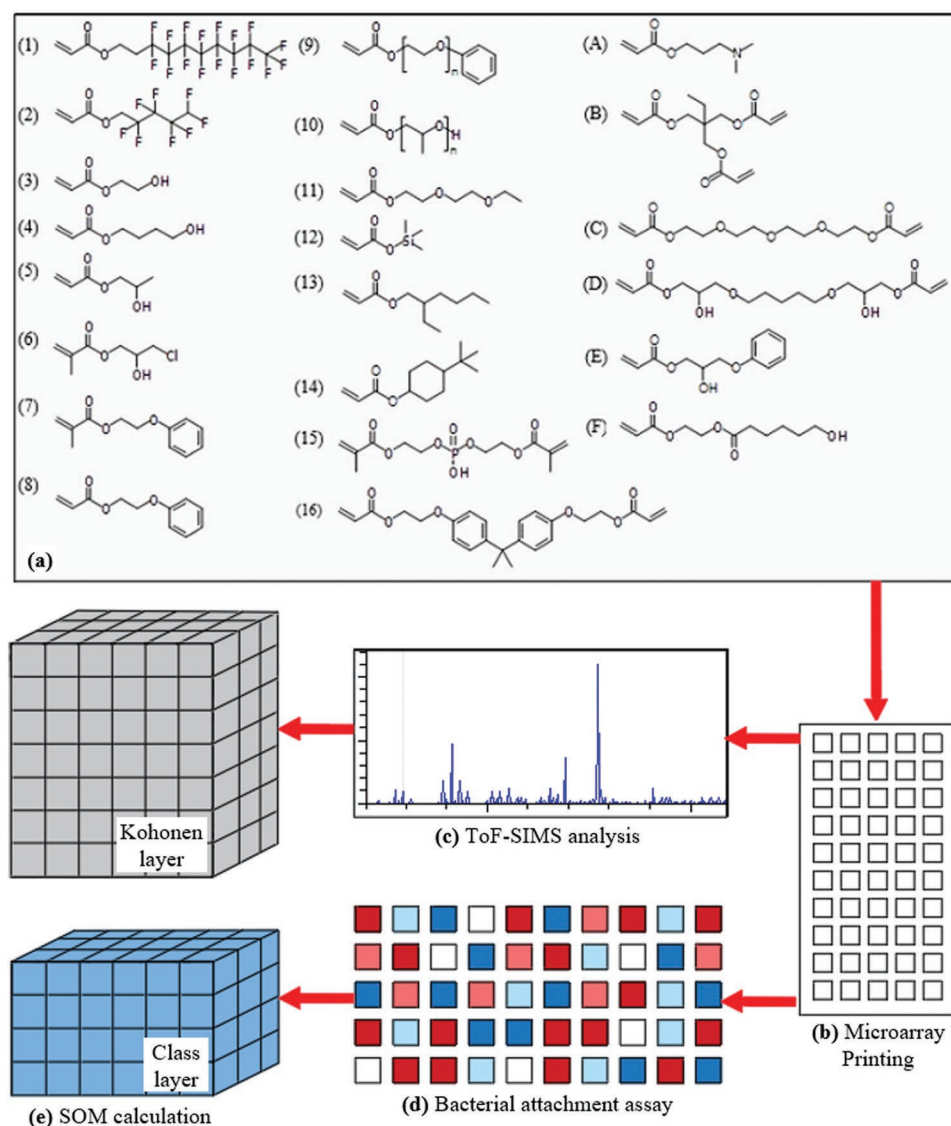


Figure 1. Schematic diagram of the workflow used to identify the best SOM model for analysis of correlation between surface chemistry and bacteria attachment properties. a) The chemical structures of the monomers, where Monomers A–F were selected as standard monomers that had previously shown low (B), medium (C,D), and high (A,E,F) bacterial attachment and tested against test Monomers 1–16 (1,2 = fluorocarbons, 3–6 = hydroxyl groups, 7,8 = comparing acrylate and methacrylate, 9–11 = comparing glycols, 12 = silyl group, 13,14 = branched versus cyclic hydrocarbon, 15,16 = comparing two diacrylates). b) Over 400 different combinations of monomers in (a) were printed on a microarray. c) ToF-SIMS analysis to create positive and negative spectra for each polymer, whereby data was acquired using peak selection. d) Intensity map after incubating the microarray with green fluorescent protein (GFP) transformed *P. aeruginosa* for 72 h. e) SOM training was done by feeding the intensities of selected peaks to the Kohonen layer and *P. aeruginosa* count to the class layer to find the relationship between surface chemistry and bacteria attachment properties.

user. Eventually, clusters of neurons with similar weight vectors would form and samples with similar ToF-SIMS intensity would be grouped into their respective clusters, hence the name self-organizing map. The term “unsupervised” indicates that no classification based on a property of interest (label) was incorporated into the UKN and its data array. The principal focus of this study was the relationship between input data and assigned class. Hence, the UKN model was not used in the current work. Its description is provided to introduce the related supervised SOMs that have been developed from it.

The CPN, also known as a pseudo-supervised model, is calculated using the same approach as the UKN. The only difference is that an associated output map is trained simultaneously as the output vector is calculated for the input signals. The unidirectional information flow is indicated by the arrow from the Kohonen layer to output layer (Figure 2a). The class membership of the neuron is therefore predicted from the class assigned to the samples and has no effect on the training of Kohonen layer. This property of the CPN makes it powerful in highlighting and confirming the indirect relationship between input data and

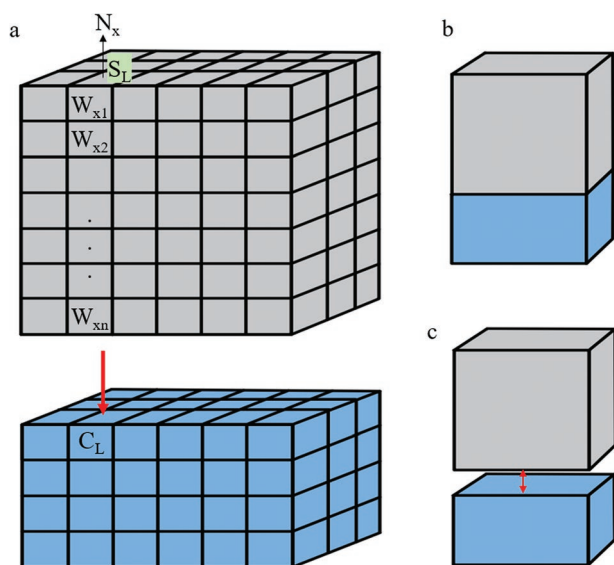


Figure 2. Schematic diagram of a) CPN, b) SKN, and c) XYF. The grey blocks represent the classical Kohonen network (UKN) and the blue blocks represent the class layers for SOM supervision. The red arrows show the direction of information flow. N represents the neuron, which is the unit element of the Kohonen map, W represents weighting for each map layer, S represent samples, and C represents the assigned class.

assigned classes. In the SKN (Figure 2b), the input data and the class vector of a sample are merged and served as input for the training of the network. After training, the combined input and output layers are decoupled and presented as a topographical map. The training of SKN is affected by the class memberships of the samples and is, therefore, a supervised model. For the XYF network, the input and output layers are calculated independently and updated simultaneously, as indicated by the arrows in Figure 2c. The common winning unit is identified by the winner shared for both input and output layers. When the training is complete, both input objects and predefined class contributed equally to the organization of the XYF map. The mathematics and strategy behind each supervised model is detailed in Melssen et al.^[27] Genetic algorithms (GA) are commonly used to identify the optimal settings for supervised SOM calculations,^[42] based on the model fitness, to avoid overfitting or underfitting. The details of the GA strategy used in this study is explained in Ballabio et al.^[43] Each chromosome consists of two sets of gene, one representing different sizes of the network (number of neurons) and the other representing numbers of training iterations (number of epochs). A chromosome with a random set of genes is used for each evaluation. Ten runs are performed for a GA calculation, with 25 evaluations for each run. Ten percent of the data is excluded for cross validation in each evaluation. Each evaluation is repeated 5 times to obtain a mean non-error rate of correctly assigned samples in the internal validation set (NER_{valid}) and a non-error rate of correctly assigned calibration sample (NER_{calib}) for the calculation of the fitness function,^[43] which is the optimization criterion for GA. Ten chromosomes with the highest fitness values for each run are filtered out as surviving units. The chromosome with the highest rate of survival and the highest fitness function is then chosen as the optimal architecture for weight analysis and prediction.

3. Experimental Section

3.1. Bacteria Attachment on a Polymer Microarray Library

Polymer microarrays were prepared using procedures described in Anderson et al. (2004);^[44] bacterial microarray screening and ToF-SIMS analyses were undertaken using the approach reported by Hook et al. (2012).^[28] In brief, stock solutions were prepared at a ratio of v/v 75% monomer, 25% dimethylformamide (DMF) and 1% w/v 2,2-dimethoxy-2-phenyl acetophenone (DMPA).^[44] Monomers were purchased from Aldrich, Scientific Polymers and Polysciences, and printed onto epoxy-coated slides (Xenopore), dip-coated into 4% w/v pHEMA (Aldrich) using 946MP6B pins (ArrayIt) and a Pixsys 5500 robot (Cartesian) or a XYZ3200 dispensing workstation (Biodot). A UV lamp (UVP Blak-Ray) was added to the workstation for polymerization by exposing the microarray to long wave UV for ≈ 10 s after each round of printing. The chemical structures of the monomers are shown in Figure 1a. Monomers 1–16 were mixed with Monomers A–F at molar ratios of 100:0 (homopolymer, 6 repeats), 90:10, 70:30, 60:40, 50:50, and 0:100 (homopolymer, 16 repeats) to create 576 solutions. The number in the name of the sample represents the molar percentage of Monomer A–F. The arrays were dried at a pressure of less than 50 mTorr for at least 7 days. Arrays were sterilized by exposure to UV for 30 min on each side, and then washed twice with phosphate buffered saline (PBS) for 30 min and then twice with medium for 30 min before use, to remove residual monomer or solvent.

The microarrays were incubated in suspensions of planktonic *P. aeruginosa* (PAO1, Nottingham strain) transformed with the constantly green fluorescent protein (GFP) expressing plasmid pME6032-GFP for 72 h.^[45] The attachment of bacteria was quantified based on the fluorescence signal (F) measured by a GenePix Autoloader 4200AL Scanner (Molecular Devices, US) with a 488 nm excitation laser and standard blue emission filter (510–560 nm). The total fluorescent intensity from the polymer spots were acquired using GenePix Pro 6 software (Molecular Devices, US). The details of bacterial growth conditions and the model for estimating bacterial count from the fluorescence signal had been previously presented by Hook et al.^[28]

3.2. ToF-SIMS Analysis of the Polymer Microarray

ToF-SIMS measurements were conducted on a ToF-SIMS IV (IONTOF GmbH, Germany) instrument operated using a monoisotopic 25 keV Bi_3^+ primary ion source in “bunched mode.” A 1 pA primary ion beam was rastered, and both positive and negative secondary ions were collected from a $100 \times 100 \mu\text{m}$ area. The typical mass resolution (at m/z 41) was approximately 6000. Charge compensation was achieved by the use of a flood gun. Positive spectra were calibrated to ions CH_3^+ , C_2H_5^+ , C_3H_7^+ and C_4H_7^+ , whilst negative spectra were calibrated to CH^- , C_2H^- , C_3H^- and C_4H^- . Identical ToF-SIMS spectra in the positive polarity were collected for sample 521 to 576, which was likely to be a sample navigation issue. These 54 samples were excluded, giving a total of 520 (N) samples for analysis.

3.3. Data Preprocessing

ToF-SIMS spectra (positive ion and negative ion) were transformed into data arrays for subsequent SOM analyses by conventional peak list formation using curated peak selections. Peaks corresponding to secondary ions that represented molecular species and structures of interest were selected for both positive and negative ion mass spectra. A total 415 peaks were selected and processed for calculation. The intensities of the selected molecular ion peaks were extracted and combined as a data array and normalized to total ion count (based on the selected peaks) per pixel. Data columns/ion images were autoscaled prior to PLS regression and range scaled independently between 0 and 1 prior to SOM training. The order of the samples was randomized and 10% of randomly selected samples were excluded as a test set, for prediction.

Five classes based on the measured bacterial fluorescence were defined for the SOM analysis of each of 520 homopolymers and copolymers, with Class 1 representing zero attachment through to Class 5 with the highest measurable attachment, as shown in Table 1.

These five classes were employed in subsequent CPN, SKN, and XYF analyses to elucidate and quantify relationships between bacterial attachment and molecular surface chemistry.

3.4. PLS Analysis

The quantified *P. aeruginosa* count was mean centered prior to PLS regression analysis, which was performed using PLS_Toolbox (Version 9.0 Eigenvector Research, Manson, WA) in MATLAB R2020a (The MathWorks Inc., USA). Contiguous blocks were used for cross validation, with 20 maximum latent variables (LVs) and 10 data folds. The number of LVs with the lowest root mean square error of cross validation (RMSECV) was selected for model calculation.

3.5. SOM Analysis

Bacterial attachment classes were used to supervise the training of three types of Kohonen networks—CPN, SKN, and XYF, with the unsupervised UKN providing a point of reference. SOMs were calculated using the Kohonen and CPANN Toolbox (Version 4.1) (Milano Chemometrics and QSAR Research Group, University of Milano-Bicocca, Italy)^[46,47] in MATLAB

Table 1. Bacterial attachment classes for *P. aeruginosa*. Samples with zero bacterial attachment were labelled as Class 1, the remaining samples being equally divided into 4 classes with increasing attachment.

Class	<i>P. aeruginosa</i> count
1	0
2	300 000–1 000 000
3	1 000 000–8 000 000
4	8 000 000–33 000 000
5	33 000 000–82 000 000

R2019b. All calculations were performed on a computing cluster (CMSS-MATLAB; 16 CPUs) at La Trobe University.

k-fold cross validation was used to optimize the SOM models. With this approach, the entire data set was partitioned into *k* distinct groups of samples, called folds. One of the *k* folds was then withheld and the remaining *k* – 1 folds are used to train the model. The performance of the model was then measured using the withheld fold. This process was repeated *k* times, such that each of the *k* folds was only withheld once. The final performance of the model was given by the mean performance across each of the *k* folds. 10 folds were used to cross-validate the models, and the venetian blinds method was used to partition the data. For the *i*th fold, where *i* ∈ {1, 2, ..., *k*}, venetian blinds includes every *k*th sample in the fold, starting from sample *i*.

The number of neurons and epochs ranging from 12–24 and 100–30 000, respectively, were used for the SOM optimization by means of GA. The bubble plots in Figure S9, Supporting Information, summarize the relative frequency of occurrence and the mean optimization criteria for each model. The optimal architecture selected for each model was used for subsequent SOMs calculations, 24 × 24 neurons and 10 000 epochs for CPN, and 22 × 22 neurons and 2000 epochs for SKN. Although the XYF map with 20 × 20 neurons and 30 000 epochs had the highest relative frequency, its mean fitness value was below average. Therefore, another set of hyperparameters (24 × 24 neurons and 5000) with a highest fitness value and second highest frequency of selection was used for XYF calculation for further comparison.

4. Results and Discussion

4.1. High Throughput Bacterial Attachment Assay

A polymer microarray was prepared containing monomers with functional groups previously shown to modulate bacterial attachment and biofilm formation.^[29,32] Sixteen monomers known to exhibit both high and low attachment were selected to ensure the inclusion of a broad range of bacterial responses (Figure 1a). These included hydrophobic (fluoro-, hydrocarbon-containing), hydrophilic (hydroxyl and ethylene glycol-containing), linear, branched, cyclic, short, and long side groups. These were mixed with 6 monomer standards that had known varied responses to bacteria. The standards were chosen from a previous microarray study as having a range of chemistries that showed high (A: 3-(dimethylamino) propyl acrylate, E: 2-hydroxy-3-phenoxypropyl acrylate, and F: 2-(acryloyloxy)ethyl 6-hydroxyhexanoate), medium (C: tetraethylene glycol diacrylate and D: [2-hydroxy-2-[4-(1-hydroxy-2-prop-2-enoyloxyethoxy)pentoxyl]ethyl] prop-2-enoate), and low (B: trimethylolpropane triacrylate) bacterial attachment.^[28] These standards were mixed with each of the 16 test monomers to form a concentration series composed of the mole ratios 1:0, 9:1, 7:3, 3:2, 1:1, and 0:1 (test:standard). In total, 406 different materials (replicates of homopolymers excluded) were UV photopolymerized by a free radical mechanism to form a combinatorial polymer microarray. After incubation with *P. aeruginosa* for 72 h, the resulting

fluorescence signals (F) obtained are shown in Figure 3a. Homopolymers of the standard Monomers A and B (shown in Figure 3a) produced low levels of bacterial attachment (Class 1 or 2), whilst for Monomers C–F, medium levels of bacterial attachment (Class 3) were observed. Most homopolymers and copolymers formed from Monomers 9–11, which contained either ethylene glycol or propylene glycol function side groups, produced high levels of bacterial attachment (Classes 4 and 5). The lowest bacterial attachment to homopolymers were observed for Monomer 13 ($0.5 \pm 0.3 \times 10^6$), whilst Monomer 1 was the most effective at maintaining low levels of bacterial attachment (Class 2) when mixed with standard Monomers C–F. The lowest overall bacterial attachment (Class 1) was observed for homopolymers of Monomer A and copolymers produced using Monomers A or B with Monomers 1, 2, or 12.

ToF-SIMS spectra collected for each sample in the microarray were coupled with the mean F value of the sample to form a new dataset and analyzed with PLS regression, a technique previously used for correlating surface chemistry and cell attachment.^[28,48,49] The R^2 value for the PLS model shown in Figure 3b was 0.70, indicating a linear relationship between

the predicted and measured bacterial attachment for this new set of polymeric microarrays. This linear relationship suggested that the attachment of *P. aeruginosa* is dependent on the surface chemistry of polymers as represented by ToF-SIMS data. The regression coefficients (RC) shown in Figure 3c represent the influence of each secondary ion on bacterial attachment. Ions with high RCs were associated with high bacterial attachment and those with low RCs were associated with high resistance to bacterial attachment. These results were consistent with those reported in Hook et al. (2012),^[28] where most oxygen-containing ions from ethylene glycol ($C_2H_2O_2^-$) and hydroxyl groups ($C_2H_6O^+$, $C_4H_9O^+$) were correlated with high bacterial attachment and hydrocarbon secondary ions representing cyclic carbon groups ($C_4H_5^-$, $C_7H_3^-$, $C_5H_5O^-$) and aliphatic groups (CH_3^+) were associated with high bacterial resistance. However, some cyclic carbon groups ($C_8H_8^+$, $C_{10}H_{12}^+$) were also found to be correlated with high bacterial attachment. Moreover, fluorine-containing (positive RC: $C_4H_8O_6F^-$, $C_2F_2^-$, negative RC: $C_2H_4O_3F_2^-$, $C_2H_2F_5^-$) and nitrogen containing ions (positive RC: $C_2H_8N^+$, negative RC: $C_3H_6NO^+$) were associated with both low and high bacterial attachment. This suggests that the relationships between sur-

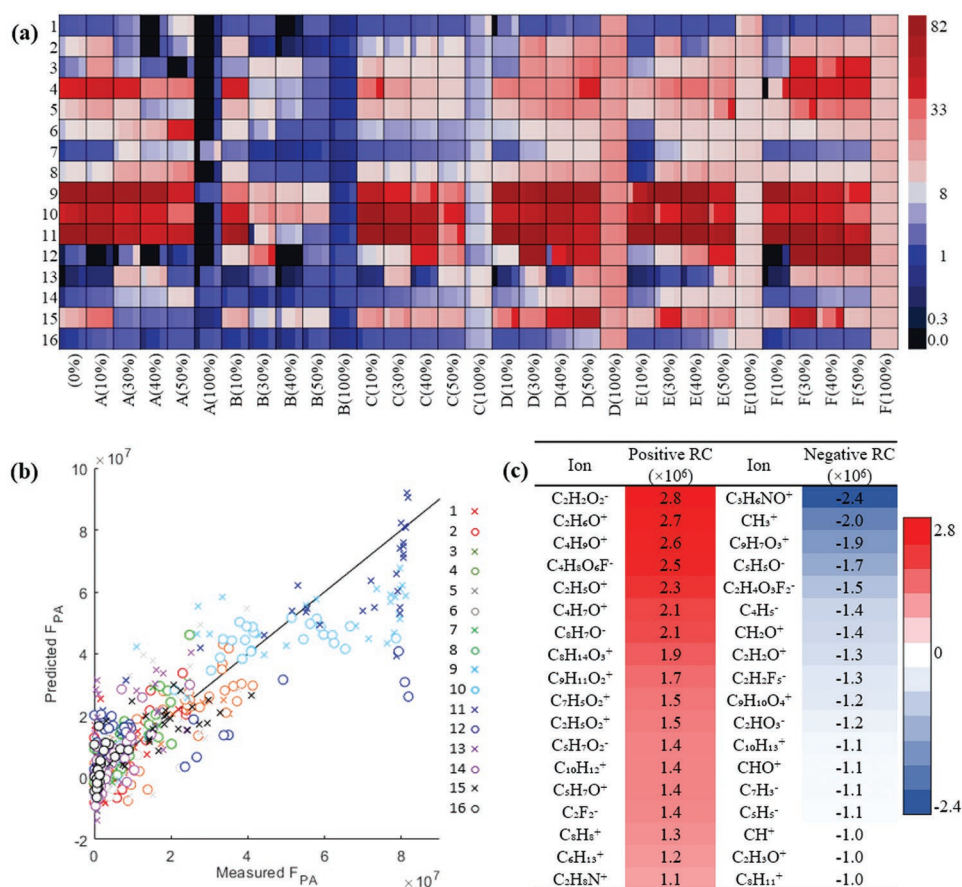


Figure 3. a) Heat map showing F , the mean fluorescence due to bacterial attachment of *P. aeruginosa* on the polymer microarray from Hook et al. (2012). The color scale is shown to the right ($\times 10^6$). Within each sample the mean (central region) \pm one standard deviation unit is shown to the left and right, respectively ($n = 3$, for 0% samples $n = 18$). b) The measured versus predicted F values of *P. aeruginosa* determined from the PLS regression model ($R^2 = 0.70$). The $y = x$ line is drawn as a guide. Polymers are grouped according to the major monomer number given in Figure 3a. c) The regression coefficient (RC) of ToF-SIMS secondary ions. Positive RCs (red) represent ions associated with high bacterial attachment and negative RCs (blue) represent ions associated with high bacterial resistance. The color scale is shown to the right ($\times 10^6$).

face chemistry and bacterial attachment are more complicated than is indicated by the linear PLS model. Given that the same fragment ion can be associated with more than one chemical structure, this is not unexpected. As such, complementary methods are needed to further explore these relationships.

4.2. Comparison of SOMs Supervised by Bacterial Attachment Data

We used SOMs to investigate bacterial–material interactions underpinning the observed bacterial attachment and compared the output weight of SOMs with the PLS model. Bacterial attachment was modulated via the surface chemistries so samples clustered by molecular composition in the SOM models should be definitively clustered by bacterial class. This approach also allowed identification of the molecular functionalities associated with low, intermediate, and high attachment. We first needed to decide which ML algorithm was best suited to

our dataset. **Figure 4** shows a comparison of supervised CPN, SKN, and XYF SOMs trained on the same ToF-SIMS data set. Note that the hyperparameters for each model were chosen based on a GA, as discussed in the Experimental section. Figure 4a–c shows SOMs trained using the five bacterial attachment classes used to supervise SOM training. An additional test of SOM supervision performance is shown in Figure 4d–f, where the class memberships were assigned randomly. We use this class scrambling to study the influence of the class labels on the clustering in the SOM.

To quantify this effect, we calculated the Moran's *I* value—a measure of spatial autocorrelation—for the SOM based on neuron class assignment. Note that the algorithm was modified to suit a toroidal topology to allow even spacing between samples. A random arrangement of classes on the SOM would produce a Moran's *I* value close to 0, while well discriminated clusters would produce a value closer to 1. We also calculated the topographical error (TE), the fraction of samples whose best and second-best winning units are not neighbors on the map, to evaluate the local

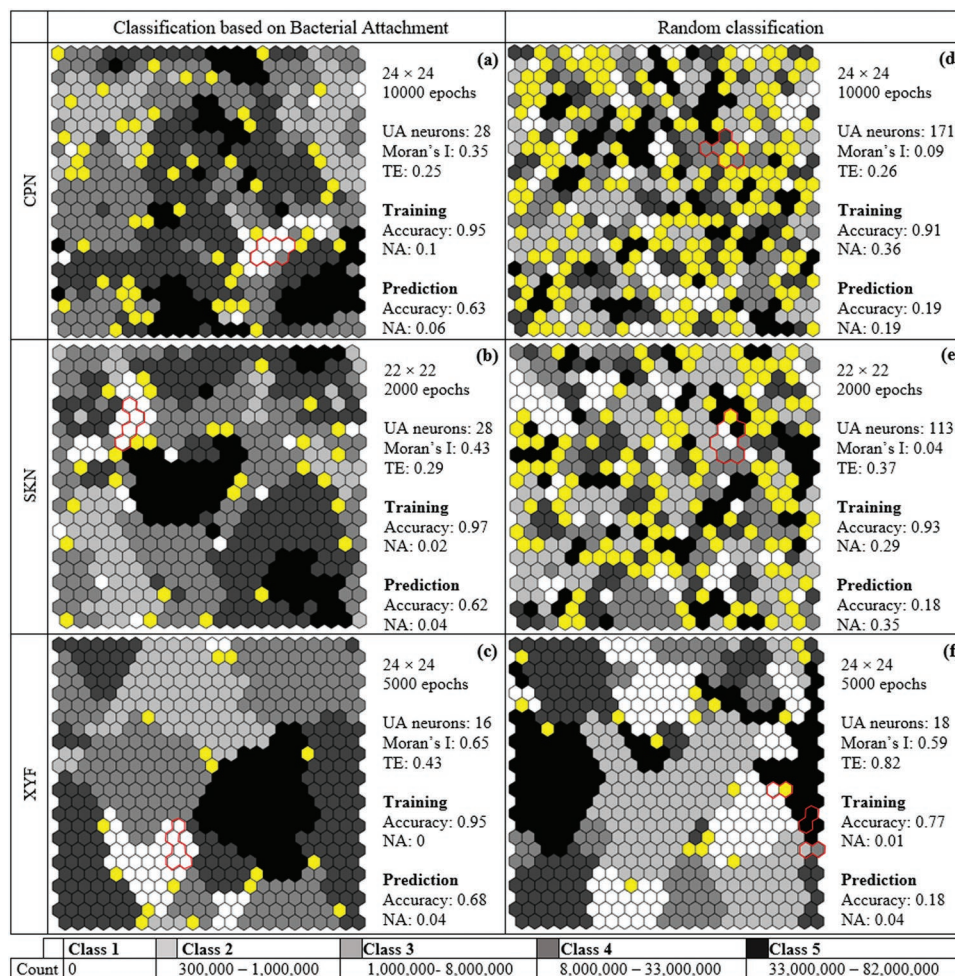


Figure 4. Topographic maps of a) CPN, b) SKN, and c) XYF with bacterial attachment as assigned class, and d) CPN, e) SKN, and f) XYF with randomly assigned class. The shade of gray represents the bacterial attachment as shown at the bottom, where the lightest shade represents Class 1 and darkest represents Class 5, and unassigned neurons are labelled yellow. The number of neurons and epochs were selected based on a genetic algorithm (GA) as shown in Figure S2, Supporting Information. The properties of each SOM, including unassigned (UA) neurons, Moran's *I*, topographic error (TE), and the accuracy and ratio not assigned sample for both training and test sets are summarized. The positions of replicates of homopolymer A are outlined in red.

map discontinuity.^[50,51] This allowed us to determine the influence of assigned class during the training of SOMs, that is, a decrease in TE associated with class assignment suggests that the class labels are decreasing the topological accuracy of the model. Finally, all SOMs were also used to predict the class assignment of the test set (10% of randomly selected samples excluded from training) to study the predictive ability of each model.

Figures 4a and 4d show CPN topographic maps where the classes are determined by bacterial attachment and by random assignment, respectively. The same TE for both CPN maps was expected, given that the class labels do not impact on model topology. To further demonstrate this, the location of replicates of homopolymer A, outlined in red, are well clustered in each case. There were significantly more unassigned neurons when a random class assignment was used. These neurons were identified as being equally associated with more than one class, and their increased abundance was due to the lack of relationship between input ToF-SIMS data and assigned class. The bacterial attachment classes (Figure 4a) form domains on the topographic map, consistent with a structure–property relationship between molecular surface composition and bacterial attachment. This resulted in a moderated Moran's *I* value of 0.35. On the contrary, the randomly assigned classes (Figure 4d) are scattered across the topographic map, as indicated by the low Moran's *I* value.

Figures 4b and 4e show SKN topographic maps with the bacterial and random class assignments, respectively. Comparing to the CPN model (Figure 4a), bacterial attachment classes formed larger, more uniform domains on the SKN map (Figure 4b), giving a higher Moran's *I* value associated with the supervision. The TE, however, was slightly higher for SKN compared to CPN, especially when random labels were used. This indicates that class assignment influenced the topological accuracy of the SOM. Interestingly, the Moran's *I* value for the SKN with random labels was lower than that of the equivalent CPN, although only marginally. Finally, the prediction accuracy was slightly higher for the SKN than the CPN, and therefore the SKN would be preferable for predictive purposes.

Figures 4c and 4f show XYF topographic maps, again with the bacterial and random class assignments, respectively. In this

case, there was a clear difference in the XYF maps, compared to both the CPN and SKN maps. For both the bacterial and random labels, very large class domains were formed on the SOM, with correspondingly high Moran's *I* values. Given that randomly assigned classes cannot have any ordered relationship with surface chemistry, this result shows a strong biasing of the outcome by the class assignment itself. This is further supported by the higher TE compared to CPN and SKN. Additionally, replicates of homopolymer A (outlined in red) are well clustered for the bacterial attachment classes (Figure 4c) but are not clustered for the randomly assigned classes (Figure 4f).

Given the Moran's *I* values and prediction accuracies for each map, particularly for the pseudo-supervised CPN, these results provide strong evidence for a relationship between surface chemistry and bacterial attachment, consistent with previous quantitative reports^[28,52] and with the PLS regression results. Although the XYF model clustered the classes well, the high TEs obtained for the XYF maps indicate that the effect of an assigned class was strongly biasing the model, leading to a loss of topological information. This is detrimental to our study, in which we sought to investigate the relationships between data topology and bacterial attachment. While the SKN provided better clustering quality than the CPN, the TE was the slightly lower for CPN, indicating higher topological accuracy. Nevertheless, given the similar results obtained with the CPN and SKN, the weights of both maps were studied to further explore the relationship between surface chemistry and bacterial attachment.

4.3. Analysis of Clusters Correlated with Bacterial Attachment

Weights from Class 1 and Class 5 neurons in the CPN and SKN, which represent polymers with lowest and highest bacterial attachment respectively, were selected for further analysis to compare the chemical groups associated with high bacterial resistance and attachment directly. **Figure 5** shows the CPN (Figure 5a) and SKN (Figure 5b) topographic maps, color coded by class weighting of Class 1 (green) and Class 5 (red). The CPN and SKN maps with complete sample labelling are

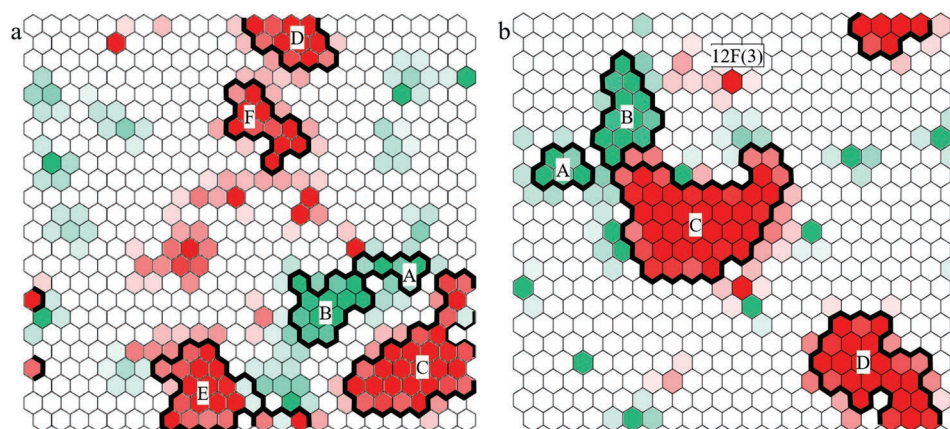


Figure 5. Topographic maps of a) CPN and b) SKN, color coded based on Class 1 (Green) and Class 5 (Red) weightings. The shades of the color are correlated to the weight of their respective class. Note that the algorithm was modified to suit a toroidal topology, hence neurons on the opposite edges are neighbors.

presented in Figures S1 and S2, Supporting Information. The CPN and SKN topographic maps illustrated in Figure 5 color coded with class weights provide insight into the clusters of the Class 1 and 5 neurons, compared to the class assignment maps shown in Figure 4. That is, the class assignment maps only show binary assignments for each class, whereas the weights show the continuous assignment fraction for each neuron. For example, although Cluster C and D in the CPN class assignment map (Figure 4a) seem to be connected, the map for Class 5 weighting (Figure 5a, red) shows that these Class 5 neurons are grouped into two sub-clusters, separated by lower weighted neurons. For both models, the smaller Class 1 clusters are labelled as A and bigger ones as B, the Class 5 clusters are then labelled alphabetically based on their distances from the Class 1 clusters.

Based on Figure 5, the SKN better separated the Class 1 clusters since Cluster A and Cluster B were separated from each other, while CPN appeared to have more clearly identified the different Class 5 clusters, as four major chemical groups were associated with high bacterial attachment instead of two in the SKN map.

The top 17 average calculated weights of the secondary ions for each cluster are summarized in Table 2, and may be used to further explore the clusters in the CPN and SKN. The secondary ions in Table 2 are color coded by their associated chemical functional groups. Note that a separate SKN (Figure S3, Supporting Information) was calculated using the replicates of the homopolymers, and the weightings of each set of replicates were analyzed to assist with assignment of the functional group. The top 15 weights for each set of homopolymers are summarized in Figure S4, Supporting Information, illustrating the chemical group assignment process.

Cluster A and Cluster B in the CPN and SKN have very similar weights, which are plotted against each other and shown in Figures S5 and S6, Supporting Information, highlighting the consistency of the SOM models. While a high weighting of nitrogen-containing secondary ions is seen for all Class 1 clusters, Cluster A, the smaller Class 1 cluster has high weighting for fluorine-containing secondary ions. This suggests that in addition to amine/nitrogen containing groups, surfaces with a combination of amine and fluorine functional groups exhibit good bacterial resistance. Figure 5 shows that SKN was more effective in separating the fluorine-containing samples from the other amine-containing materials. A selection of the major secondary ions associated with high bacterial attachment and resistance are illustrated in Figure 6 for improved visualization. Some of the secondary ions such as $\text{C}_2\text{H}_4\text{O}^+$ and $\text{C}_2\text{H}_4\text{O}_2^+$ may be derived from glycol and hydroxyl groups, which could complicate conventional linear multivariate analysis.

For the CPN Class 5 clusters, Cluster C consisted mostly of glycol containing secondary ions, Cluster D mostly phenyl, Cluster E mostly hydroxyl, and Cluster F mostly cyclic moieties. For SKN, these four chemical groups merged into two major clusters, where Cluster C contained a high proportion of glycol and hydroxyl groups, and Cluster D contained mostly phenyl and other cyclic groups. The ability of SKN to identify a contaminated sample is demonstrated by the isolated Class 5 neuron containing sample 12F(3). High levels of nitrogen- and phenyl-containing secondary ions suggest that 12F(3) is contaminated,

since Monomers 12 and F did not contain amine and phenyl groups. The nitrogen and phenyl contamination may have been from residual DMF ($\text{C}_3\text{H}_7\text{NO}$) and DMPA ($\text{C}_{16}\text{H}_{16}\text{O}_3$), respectively, or inadequate washing of the microarray pins during generation of the polymer spots and potentially crosstalk from the washing procedures. The sample still identified as Class 5 although nitrogen-containing secondary ions was previously shown to be associated with negligible bacterial attachment.

From Table 2, it appears that the CPN more clearly separated different functional groups associated with bacterial attachment, since the phenyl containing samples were separated from glycol and hydroxyl moieties. However, among the phenyl containing monomers, only Monomer 9 was found in the Class 5 clusters, whereas Monomers 7 and 8 also contained a phenyl group. Most of the samples located in the Class 5 clusters contain Monomers 9, 10 and 11, suggesting that the major chemical group contributing to high bacterial attachment was the long glycol chain, not the phenyl group of Monomer 9. Hence, the SKN appeared to more clearly identify the main functional groups associated with high bacterial attachment ability over CPN because of the supervised learning process.

4.4. Using the SOM to Identify Key Interaction Variables

It is valuable to compare the PLS regression and SOM results directly, to identify both similar and conflicting outcomes. First, we note that both methods identified $\text{C}_2\text{H}_2\text{O}_2^-$ (ethylene glycol), $\text{C}_4\text{H}_9\text{O}^+$ (hydroxyl) and C_8H_8^+ (cyclic) to be associated with bacterial attachment and CH^+ (aliphatic), $\text{C}_3\text{H}_3\text{O}_2\text{F}_2^-$ and $\text{C}_2\text{H}_2\text{F}_5^-$ to be associated with high bacterial resistance. However, it is interesting that PLS regression did not identify some of the fragment ions associated with high or low bacterial attachment, such as $\text{C}_5\text{H}_8\text{O}_2^+$, $\text{C}_4\text{H}_5\text{O}_2^+$ and ions containing amine groups. One possible reason is that any given fragment ion may not be unique to a specific polymer, and hence may not correlate strongly (positively or negatively) with bacterial attachment. This is critical, as it indicates a limitation of the PLS regression approach, which seeks to identify correlations between individual fragment ions and bacterial attachment. The SOM methods, on the other hand, were used to explore the relationship between data topology and bacterial attachment, and therefore reveal complementary information to PLS. In particular, the SOM identified clusters of high or low bacterial attachment surface chemistries, from which combinations of fragment ions associated with each cluster could be studied and compared.

As an additional validation of the SOM results, and to show how PLS regression and the SOM can be used together, we calculated a second PLS regression model, this time including additional interaction variables. These interaction variables were calculated using the 3 top weighted positive and negative fragment ions associated with each of the Class 1 and Class 5 SKN clusters. The intensities for each positive ion were multiplied with those from each negative ion within the same cluster, giving nine additional interaction variables for each cluster (totaling 36 interaction variables across the 4 clusters).

We compared the two PLS regression models to investigate whether the additional interaction variables improved the model. A slightly higher R^2 (Figure S7a, Supporting

CPN																				
Class 1					Class 5															
Cluster A			Cluster B		Cluster C			Cluster D		Cluster E		Cluster F								
1, A, F			A, 1, 12		9, 10, 11, A, C, D, E, F			9, 11, C, D, E, F		10, 11, A, B, C, D, E, F		3, 4, 9, 11, 15, D, F								
Pos SE	Weight	Neg SE	Weight	Neg SE	Pos SE	Weight	Neg SE	Pos SE	Weight	Neg SE	Pos SE	Weight	Neg SE							
$C_6H_8N_2^+$	0.88	F^-	0.62	$C_4H_8N^+$	0.86	$C_2H_4O_2^+$	CF^-	0.66	$C_3H_{11}^+$	0.50	C_2HO^-	0.70	$C_3H_5^+$	0.62	$C_3H_2O_2^-$	0.68	$C_4H_7^+$	0.50	C_4HO^-	0.68
$C_6H_8N^+$	0.79	CF^-	0.62	$C_3H_9N^+$	0.85	$C_2H_4O^+$	C_2HO^-	0.49	$C_6H_8^+$	0.46	C_4HO^-	0.67	$C_3H_7O^+$	0.61	$C_4H_2O_2^-$	0.64	$C_3H_7^+$	0.48	C_3HO^-	0.63
$C_5H_6N^+$	0.72	$C_3H_3O_2F_2^-$	0.60	CH_4N^+	0.84	$C_3H_2^+$	F^-	0.48	$C_8H_8^+$	0.44	$C_4H_2O_2^-$	0.64	$C_3H_3O_2^+$	0.57	C_2HO^-	0.60	$C_7H_9^+$	0.47	$C_3H_2O^-$	0.62
$C_4H_4N^+$	0.72	C_3N^-	0.59	$C_4H_9N^+$	0.80	$C_2H_5O_2^+$	$C_4H_2O_2^-$	0.44	$C_3H_7^+$	0.43	$C_6H_3^-$	0.61	$C_3H_5O^+$	0.55	$C_3H_2O^-$	0.57	$C_6H_7^+$	0.47	$C_4H_2O_2^-$	0.61
$C_4H_4NO^+$	0.68	$C_4F_3^-$	0.59	$C_4H_6N^+$	0.80	$C_3H_8O_2^+$	$C_3H_3O^-$	0.44	$C_3H_5^+$	0.41	$C_4H_3^-$	0.60	$C_3H_8O^+$	0.54	$C_2H_3O^-$	0.55	$C_6H_8^+$	0.46	$C_4H_3O_2^-$	0.59
$C_6H_{10}N^+$	0.68	C_2F^-	0.58	$C_3H_8N^+$	0.79	$C_2H_6O^+$	$C_3H_2O_2^-$	0.43	$C_3H_9^+$	0.40	$C_3H_2O^-$	0.59	$C_3H_4O^+$	0.53	CHO^-	0.54	$C_3H_5^+$	0.46	$C_3H_2O^-$	0.59
$C_3H_2N^+$	0.65	HF_2^-	0.56	$C_3H_6N^+$	0.79	$C_2H_3O^+$	$C_3H_2O^-$	0.43	$C_2H_5^+$	0.39	C_3HO^-	0.59	$C_4H_3O_2^+$	0.52	$C_3O_2^-$	0.53	$C_3H_7O^+$	0.43	$C_3H_3O_2^-$	0.58
$C_4H_6N^+$	0.64	$C_3H_4O^-$	0.56	$C_6H_8N_2^+$	0.78	CHO_2^+	$C_3O_2^-$	0.44	$C_3H_8O_2^+$	0.37	$C_3H_3^-$	0.59	$C_3H_9O^+$	0.52	$C_3H_2O^-$	0.53	$C_3H_8O_2^+$	0.40	$C_6H_4^-$	0.58
$C_3H_3O^+$	0.63	CF_3^-	0.52	$C_6H_8N^+$	0.77	$C_3H_7O^+$	$C_3H_4^-$	0.44	$C_4H_6^+$	0.37	$C_3H_3O_2^-$	0.58	$C_3H_8O_2^+$	0.49	$C_2H_2O_2^-$	0.52	$C_3H_9^+$	0.38	$C_3H_3O_2^-$	0.58
$C_3H_4NO_2^+$	0.62	C_4HN^-	0.51	$C_6H_{10}N^+$	0.73	CH_3O^+	$C_4H_6NO^-$	0.43	$C_6H_8^+$	0.36	$C_3H_4^-$	0.57	CH_4O^+	0.48	C_3HO^-	0.52	$C_4H_6^+$	0.36	$C_3H_2O_2^-$	0.58
CH_4N^+	0.61	$C_3F_3^-$	0.50	$C_3H_7N^+$	0.73	$C_2H_2O^+$	$C_3H_2O^-$	0.41	$C_4H_9^+$	0.36	$C_3H_3O^-$	0.57	$C_3H_6O^+$	0.48	$C_2HO_2^-$	0.51	$C_{10}H_{10}^+$	0.36	C_3HO^-	0.58
$C_3H_7NF_2^+$	0.61	$C_2H_4O_3F_2^-$	0.49	$C_3H_{11}N^+$	0.71	$C_3H_5^+$	CHO^-	0.41	$C_7H_3O_2^+$	0.36	C_6HO^-	0.56	$C_4H_2O_2^+$	0.45	$C_3H_3O_2^-$	0.47	$C_4H_9^+$	0.36	$C_6H_3^-$	0.57
CH^+	0.61	$C_2H_2F_3^-$	0.49	$C_3H_4N^+$	0.70	$C_3H_5^+$	$C_3H_2O_2^-$	0.40	$C_6H_7^+$	0.35	$C_6H_4^-$	0.54	$C_3H_3O_2^+$	0.45	$C_3HO_3^-$	0.47	$C_8H_8^+$	0.36	C_4O^-	0.56
$C_3H_2^+$	0.59	C_3F^-	0.48	$C_2H_5N^+$	0.64	CH_2^+	$C_2H_3O^-$	0.40	$C_3H_7^+$	0.35	$C_4H_2O^-$	0.52	$C_3H_3O_2^+$	0.45	$C_3H_3O_2^-$	0.47	$C_9H_9O_2^+$	0.35	$C_3H_3^-$	0.55
CH_2^+	0.59	C_4F^- </																		

[illegible]

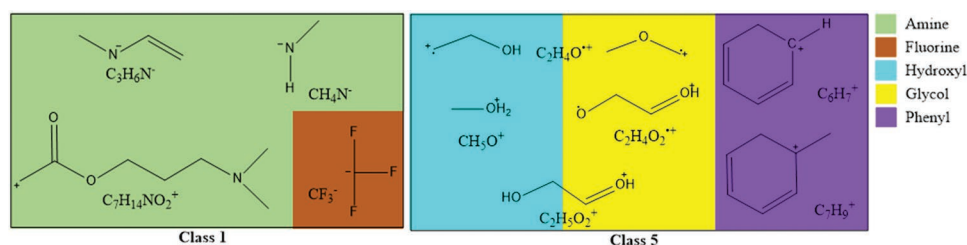


Figure 6. Schematic diagram of the secondary ions associated with high bacterial resistance (left) and high bacterial attachment (right). Secondary ions are color coded according to their functional groups.

Information) value and lower root mean square error of cross validation (RMSECV) (Figure S7b, Supporting Information) were obtained after adding the interaction variables, indicating that the interaction variables slightly improved the prediction accuracy of the PLS model. Interaction variables with either high or low correlation with bacterial attachment were also identified. **Table 3** shows the 6 most highly positively and negatively correlating interaction variables (the complete RCs calculated for all 36 interaction variables are shown in Table S8, Supporting Information). Also shown are the RCs associated with each individual ion from the interaction as a comparison, as calculated from the original PLS regression model without interaction variables.

There are several important observations relating to Table 3. In all cases, the interaction variables exhibited higher RCs than their constituent fragment ions alone. Interaction variables containing cyclic hydrocarbon coupled with oxygen-containing ions were most positively correlated with bacterial attachment. This agrees with the outcome of SKN that indicated that a phenyl group attached to a long glycol chain (Monomer 9) had a much stronger bacterial attachment ability compared to phenyl group alone (Monomers 7 and 8). Importantly, the $C_4H_2O_2^-$, C_2HO^- and C_4HO^- ions were not individually strongly positively correlated to attachment, instead exhibiting low positive or even negative correlation. This suggests that these ions are not generally correlated with bacterial attachment, however when present

with cyclic hydrocarbon ions are strongly correlated. Results in Table 3 also help to explain why positive RCs were obtained for some cyclic groups in the original PLS model in Figure 3, despite the fact that cyclic groups were previously found to be associated with high bacterial resistance.^[28]

The interaction variables featuring ions containing nitrogen and fluorine further emphasize the limitations of the original PLS regression model. Individually, both $C_4H_4N^+$ and F^- were found to be positively correlated with bacterial attachment. However, the interaction of the two ions gave a highly negative RC. In fact, all interaction terms for nitrogen- and fluorine-containing ions were found to be correlated with bacterial resistance (Table S9, Supporting Information). This consistent result further emphasizes that fluorine groups coupled with amine groups (as identified by the SKN) lead to high *P. aeruginosa* resistance. More studies need to be done to fully understand why the combination of these two functional groups prevents *P. aeruginosa* attachment. A combination of several factors, such as hydrophobicity, chemical reaction between polymer and bacterial surfaces, and orientation and arrangement of the functional groups, could be contributing to this observation.

Together, these results highlight the importance of complementing standard PLS regression analysis with techniques such as the SOM, which can provide further information about potential interactions between fragment ions. In this regard, understanding data topology is critical to accurately interpret and improve upon the information provided by PLS.

Table 3. Regression coefficients of interaction variables selected from the SKN clusters.

Cluster	Interaction variable	RC	Pos ion RC ($\times 10^5$)	Neg ion RC ($\times 10^5$)	$\times 10^5$
D	$C_8H_8^+ \times C_4H_2O_2^-$	20.1	12.9	1.2	20
D	$C_8H_8^+ \times C_2HO^-$	18.6	12.9	4.3	16
D	$C_8H_8^+ \times C_4HO^-$	10.8	12.9	-3.2	12
D	$C_2H_4O_2^+ \times C_2HO^-$	8.2	5.6	4.3	8
D	$C_7H_9^+ \times C_4H_2O_2^-$	6.8	-1.1	1.2	4
C	$C_2H_4O_2^+ \times CF^-$	5.9	5.6	5.3	0
A	$C_4H_4N^+ \times F^-$	-5.1	1.1	2.4	-2
A	$C_6H_8N^+ \times F^-$	-5.2	-0.5	2.4	-4
A	$C_4H_4N^+ \times C_3H_3O_2F_2^-$	-5.6	1.1	-9.1	-6
A	$C_6H_8N_2^+ \times F^-$	-6.3	-5.3	2.4	-8
A	$C_6H_8N^+ \times C_3H_3O_2F_2^-$	-10.6	-0.5	-9.1	-10
A	$C_6H_8N_2^+ \times C_3H_3O_2F_2^-$	-12.7	-5.3	-9.1	-12

5. Conclusion

The relationship between ToF-SIMS spectra and *P. aeruginosa* attachment for a set of polymer combinations previously found to be associated with low, moderate, and high resistance to bacterial attachment was assessed by PLS and SOM models. The observed correlation between surface chemical data and bacterial attachment suggests that for the polymeric materials studied, the surface chemistry plays a key determining role in the bacterial response.

We explored different SOM variants and their ability to reveal information about structure–property relationships in a polymer microarray. In our case, the use of class labels biased the ability of the XYF to learn the data topology, undesirable for identifying the relationship between polymer surface chemistry and performance. For the SKN, while class membership clearly influenced topographic map formation, the modelled topology much more closely resembled that of the pseudo-supervised

CPN, suggesting lower influence of the class labels and therefore a higher influence of the ToF-SIMS molecular data.

Comparing the weights of clusters representing the highest and lowest bacterial attachment for CPN and SKN suggested that SKN had a stronger ability to cluster and identify major chemical functional groups associated with bacterial resistance and bacterial attachment. The weights of the SKN suggested that fluorine- and nitrogen-containing groups hinder bacterial attachment while hydroxyl and long glycol groups promote bacterial attachment. The weights also showed that the contaminated samples or other outliers could be identified in the SKN topographic map as isolated neurons. This powerful ability of the SKN to identify contaminated samples allows them and other outliers, which could complicate conventional regression analyses, to be eliminated with justification.

While PLS is a powerful tool to detect the correlation of an individual ion, SOMs successfully detected ions associated low and high *P. aeruginosa* resistance that were not easily detected by the PLS model. Moreover, SOMs were able to identify different classes of materials with different chemistry that behaved similarly in terms of bacteria-surface interactions. Combining a SOM with PLS showed that interaction variables are important and highlight limitations of the original PLS. Coupling these two data analysis models is therefore crucial for a much deeper understanding of the complex interactions between surface chemistry and bacterial surfaces. By coupling the two models, further analysis could be done with different types of bacteria to study the universality of the bacterial resistance property of the polymer array.

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Acknowledgements

This work was supported by an Office of National Intelligence – National Intelligence and Security Discovery Research Grant (NI210100127) funded by the Australian Government. This work was performed in part at the S (ANFF), a company established under the National Collaborative Research Infrastructure Strategy, through the La Trobe University Centre for Materials and Surface Science. The authors acknowledge the Milano Chemometrics and QSAR Research Group for the development of the Kohonen and CP-ANN Toolbox for MATLAB.^[46,47] Y.M. is supported by NSF Award #OIA-1655740. Jing Yang is kindly acknowledged for assistance with preparation of microarrays. David Scurr is kindly acknowledged for his assistance in acquiring the ToF-SIMS spectra.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Keywords

artificial neural networks, bacterial attachment, mass segmentation, microarrays, polymers, self-organizing maps, ToF-SIMS

Received: November 8, 2022

Revised: January 9, 2023

Published online:

- [1] D. Davies, *Nat. Rev. Drug Discovery* **2003**, 2, 114.
- [2] M. Cámara, W. Green, C. E. MacPhee, P. D. Rakowska, R. Raval, M. C. Richardson, J. Slater-Jefferies, K. Steventon, J. S. Webb, *NPJ Biofilms Microbiomes* **2022**, 8, 42.
- [3] L. E. Fisher, A. L. Hook, W. Ashraf, A. Yousef, D. A. Barrett, D. J. Scurr, X. Chen, E. F. Smith, M. Fay, C. D. J. Parmenter, R. Parkinson, R. Bayston, *J. Controlled Release* **2015**, 202, 57.
- [4] R. O. Darouiche, I. I. Raad, S. O. Heard, J. I. Thornby, O. C. Wenker, A. Gabrielli, J. Berg, N. Khardori, H. Hanna, R. Hachem, R. L. Harris, G. Mayhall, *N. Engl. J. Med.* **1999**, 340, 1.
- [5] I. Raad, R. Darouiche, J. Dupuis, D. AbiSaid, A. Gabrielli, R. Hachem, M. Wall, R. Harris, J. Jones, A. Buzaid, C. Robertson, S. Shenaq, P. Curling, T. Burke, C. Ericsson, S. Greenberg, N. Hanania, D. Yosher, D. Gibson, M. Reardon, P. Reardon, T. Darnule, M. Mansouri, K. Rolston, E. Whimbey, C. Bivins, A. Huaranga, K. Price, H. Safar, *Ann. Intern. Med.* **1997**, 127, 267.
- [6] K. Yorganci, C. Krepel, J. A. Weigelt, C. E. Edmiston, *Intensive Care Med.* **2002**, 28, 438.
- [7] K. Jaeger, A. Osthaus, J. Heine, H. Ruschulte, C. Kuhlmann, H. Weissbrodt, A. Ganser, M. Karthaus, *Chemotherapy* **2001**, 47, 50.
- [8] J. I. Greenfeld, L. Sampath, S. J. Popilskis, S. R. Brunnert, S. Stylianos, S. Modak, *Crit. Care Med.* **1995**, 23, 894.
- [9] D. R. Guay, *Drugs* **2001**, 61, 353.
- [10] L. Caillier, E. T. de Givenchy, R. Levy, Y. Vandenberghe, S. Geribaldi, F. Guittard, *Eur. J. Med. Chem.* **2009**, 44, 3201.
- [11] R. O. Darouiche, M. D. Mansouri, P. V. Gawande, S. Madhyastha, *J. Antimicrob. Chemother.* **2008**, 61, 651.
- [12] H. Koo, R. N. Allan, R. P. Howlin, P. Stoodley, L. Hall-Stoodley, *Nat. Rev. Microbiol.* **2017**, 15, 740.
- [13] M. Balain, S. J. Oddie, W. McGuire, *Cochrane Database Syst. Rev.* **2015**, CD011078.
- [14] N. M. Lai, N. Chaiyakunapruk, N. A. Lai, E. O'Riordan, W. S. C. Pau, S. Saint, *Cochrane Database Syst. Rev.* **2016**, CD007878.
- [15] K. Schumm, T. B. L. Lam, *Cochrane Database Syst. Rev.* **2008**, 2, CD004013.
- [16] G. Tokmaji, H. Vermeulen, M. C. A. Müller, P. H. S. Kwakman, M. J. Schultz, S. A. J. Zaat, *Cochrane Database Syst. Rev.* **2015**, CD009201.
- [17] R. D. Monds, G. A. O'Toole, *Trends Microbiol.* **2009**, 17, 73.
- [18] B. D. Ratner, A. Chilkoti, D. G. Castner, *Clin. Mater.* **1992**, 11, 25.
- [19] B. D. Ratner, B. J. Tyler, A. Chilkoti, *Clin. Mater.* **1993**, 13, 71.
- [20] D. J. Graham, D. G. Castner, *Biointerphases* **2012**, 7, 49.
- [21] M. S. Wagner, D. G. Castner, *Langmuir* **2001**, 17, 4649.
- [22] A. L. Hook, D. J. Scurr, *Surf. Interface Anal.* **2016**, 48, 226.
- [23] A. D. Celiz, H. C. Harrington, A. L. Hook, *Appl. Surf. Sci.* **2014**, 313, 926.
- [24] A. J. Urquhart, M. Taylor, D. G. Anderson, R. Langer, M. C. Davies, M. R. Alexander, *Anal. Chem.* **2008**, 80, 135.

- [25] P. J. Cumpson, I. W. Fletcher, N. Sano, A. J. Barlow, *Surf. Interface Anal.* **2016**, 48, 1328.
- [26] P. J. Cumpson, N. Sano, I. W. Fletcher, J. F. Portoles, M. Bravo-Sanchez, A. J. Barlow, *Surf. Interface Anal.* **2015**, 47, 986.
- [27] W. Melssen, R. Wehrens, L. Buydens, *Chemom. Intell. Lab. Syst.* **2006**, 83, 99.
- [28] A. L. Hook, C. Y. Chang, J. Yang, J. Luckett, A. Cockayne, S. Atkinson, Y. Mei, R. Bayston, D. J. Irvine, R. Langer, D. G. Anderson, P. Williams, M. C. Davies, M. R. Alexander, *Nat. Biotechnol.* **2012**, 30, 868.
- [29] O. Sanni, C. Chang, D. G. Anderson, R. Langer, M. C. Davies, P. M. Williams, P. Williams, M. R. Alexander, A. L. Hook, *Adv. Healthcare Mater.* **2015**, 4, 695.
- [30] P. Mikulskis, A. Hook, A. A. Dundas, D. Irvine, O. Sanni, D. Anderson, R. Langer, M. R. Alexander, P. Williams, D. A. Winkler, *ACS Appl. Mater. Interfaces* **2018**, 10, 139.
- [31] P. Mikulskis, M. R. Alexander, D. A. Winkler, *Adv. Intell. Syst.* **2019**, 1, 1900045.
- [32] V. C. Epa, A. L. Hook, C. Chang, J. Yang, R. Langer, D. G. Anderson, P. Williams, M. C. Davies, M. R. Alexander, D. A. Winkler, *Adv. Funct. Mater.* **2014**, 24, 2085.
- [33] T. Kohonen, in *Self-Organization and Associative Memory*, Springer-Verlag, Berlin, Heidelberg, Germany **1989**, Vol. 8.
- [34] W. S. McCulloch, W. Pitts, *Bull. Math. Biol.* **1990**, 52, 99.
- [35] Y.-Y. Chen, Y.-H. Lin, C.-C. Kung, M.-H. Chung, I. H. Yen, *Sensors* **2019**, 19, 2047.
- [36] T. Kohonen, *Biol. Cybern.* **1982**, 43, 59.
- [37] M. Liukkonen, M. Heikkinen, T. Hiltunen, E. Hälikkää, R. Kuivalainen, Y. Hiltunen, *Energy* **2011**, 36, 339.
- [38] S. H. Romanić, G. Vuković, D. Klinčić, D. Antanasijević, *Sci. Total Environ.* **2018**, 628–629, 198.
- [39] W. Gardner, S. M. Cutts, B. W. Muir, R. T. Jones, P. J. Pigram, *Anal. Chem.* **2019**, 91, 13855.
- [40] P. Louis, A. Seret, B. Baesens, *World Dev.* **2013**, 46, 197.
- [41] W. Gardner, A. L. Hook, M. R. Alexander, D. Ballabio, S. M. Cutts, B. W. Muir, P. J. Pigram, *Anal. Chem.* **2020**, 92, 6587.
- [42] Z. Ma, *A Genetic SOM Clustering Algorithm for Intrusion Detection* (Ed: J. Wang, X. F. Liao, Z. Yi), Springer, Berlin, Heidelberg **2005**.
- [43] D. Ballabio, M. Vasighi, V. Consonni, M. Kompany-Zareh, *Chemom. Intell. Lab. Syst.* **2011**, 105, 56.
- [44] D. G. Anderson, S. Levenberg, R. Langer, *Nat. Biotechnol.* **2004**, 22, 863.
- [45] S. N. A. Qazi, C. E. D. Rees, K. H. Mellits, P. J. Hill, *Microb. Ecol.* **2001**, 41, 301.
- [46] D. Ballabio, V. Consonni, R. Todeschini, *Chemom. Intell. Lab. Syst.* **2009**, 98, 115.
- [47] D. Ballabio, M. Vasighi, *Chemom. Intell. Lab. Syst.* **2012**, 118, 24.
- [48] Y. Mei, K. Saha, S. R. Bogatyrev, J. Yang, A. L. Hook, Z. I. Kalcioğlu, S.-W. Cho, M. Mitalipova, N. Pyzocha, F. Rojas, K. J. Van Vliet, M. C. Davies, M. R. Alexander, R. Langer, R. Jaenisch, D. G. Anderson, *Nat. Mater.* **2010**, 9, 768.
- [49] J. Yang, Y. Mei, A. L. Hook, M. Taylor, A. J. Urquhart, S. R. Bogatyrev, R. Langer, D. G. Anderson, M. C. Davies, M. R. Alexander, *Biomaterials* **2010**, 31, 8827.
- [50] F. Forest, M. Lebbah, H. Azzag, J. Lacaille, *arXiv preprint arXiv* **2020**, <https://doi.org/10.48550/arXiv.2011.05847>.
- [51] G. T. Breard, *Degree Thesis*, University of Rhode Island **2017**.
- [52] A. L. Hook, D. G. Anderson, R. Langer, P. Williams, M. C. Davies, M. R. Alexander, *Biomaterials* **2010**, 31, 187.