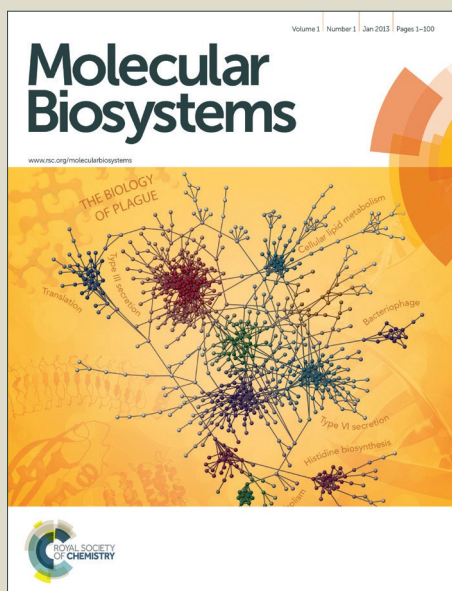


# Molecular BioSystems

Accepted Manuscript



This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

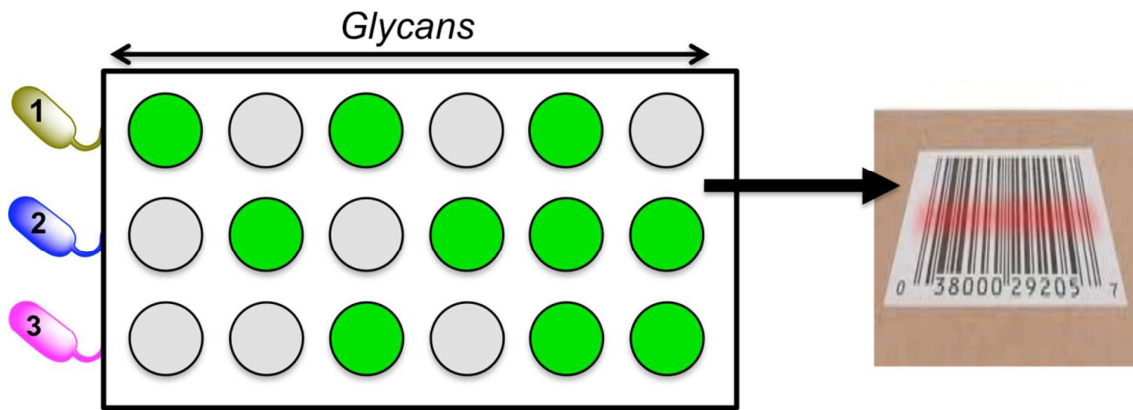
*Accepted Manuscripts* are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

You can find more information about *Accepted Manuscripts* in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.



[www.rsc.org/molecularbiosystems](http://www.rsc.org/molecularbiosystems)



'Barcoding' how bacteria bind to simple sugar-surfaces using a training algorithm is used as a powerful identification tool.



Journal Name

COMMUNICATION

## Discrimination between Bacterial Species by Ratiometric analysis of their Carbohydrate Binding Profile

Received 00th January 20xx,  
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

www.rsc.org/

Lucienne Otten<sup>a</sup>, Elizabeth Fullam<sup>b</sup> and Matthew I. Gibson<sup>a,\*</sup>

**Antibiotic resistance is a global health concern meaning there is an urgent need for new treatments and diagnostics. Here glycosylated surfaces are used to profile the binding patterns of a range of Gram-negative, Gram-positive and Mycobacteria. This enables the creation of 'barcodes' to enable identification and discrimination between the strains, which could not be achieved by single-point glycan binding and offers a new concept in bacteria detection.**

Bacteria cause diseases including meningitis, pneumonia, diarrhea and many hospital acquired infections such as *Staphylococcus aureus*<sup>1</sup> or *Clostridium difficile*<sup>2</sup> and are responsible for millions of deaths every year. 'Old' diseases such as tuberculosis, caused by the bacterium *Mycobacterium tuberculosis*, for which antibiotics had previously proven to be effective, are once again on the rise<sup>3-5</sup> with an estimated 2 billion carriers globally. Since penicillin was discovered in 1929 it revolutionized the treatment of bacterial infections<sup>6</sup> but the current drug pipeline has dried up alongside development of resistance.<sup>7</sup> The spread of resistance is very rapid; for example methicillin was clinically used in 1960 with the first case of methicillin resistant *S. aureus* reported in 1961.<sup>8</sup> These factors have contributed to the prediction that by 2050 the number of deaths associated with antimicrobial resistance will outnumber those caused by cancer.<sup>9,10</sup> The problem of antibiotic resistance is propagated by the lack of rapid point-of-care diagnostics tools to ensure correct use of the remaining drugs. Most of the current Food and Drug Administration (FDA) approved detection methods rely on a culturing step; the bacteria are isolated and grown for visual, or other examination methods, which can take several days<sup>11</sup> and some strains are intrinsically challenging to culture. Since Robert Koch first described culturing for identification of bacteria over 140 years ago the techniques used has hardly changed since.<sup>12,</sup>

<sup>13</sup> However biofilm based infections are now becoming the most common problem, with these bacteria proving incredibly challenging to culture.<sup>14</sup> It is estimated that only 1% of bacteria are culturable on standard culture media and several of the species that commonly cause hospital acquired infections can enter a viable but not culturable phase where they are not detectable through culturing techniques.<sup>12, 14, 15</sup> The emerging alternative is the use of polymerase chain reaction (PCR) or sequencing methods. PCR based processes include; (i) the analysis of the full genome of microbial species and identification through comparison with a database<sup>16, 17</sup> and (ii) Ibis, a method which involves selective PCR of a species-specific element of DNA which is flanked by highly conserved regions. Whilst PCR based methods are very rapid (with Ibis reportedly only taking 6 hours) they are still relatively expensive (Ibis needs a mass spectrometer) and rely on extraction of nucleic acids from within bacteria, which can be problematic in Gram-positive species due to their thick peptidoglycan wall, which is difficult to lyse.<sup>14</sup>

Many bacteria infections require an initial adhesion step, for example, the high mannose-binding adhesin FimH is a crucial virulence factor found in uropathogenic *Escherichia coli* and among enterobacteriaceae.<sup>18,19</sup> Biofilm based infections are responsible for many chronic bacteria associated conditions such as pneumonia<sup>20</sup> and tuberculosis<sup>21</sup> and also the most difficult to culture.<sup>22-24</sup> Considering the above, adhesins are potential targets for novel prophylactic anti-adhesion therapies<sup>25, 26</sup> or as diagnostics.<sup>27-30</sup> Several nanoparticle-based detections systems exploiting adhesins have been reported including galabiose particles for the detection of *Streptococcus suis*.<sup>20, 31</sup> Mannose conjugated gold nanoparticles have been shown to detect FimH positive strains of *E. coli* in a colorimetric assay.<sup>27</sup> These particle-based systems highlight the potential for carbohydrates in detection but are intrinsically limited by the promiscuity of glycan-lectin (carbohydrate binding proteins) interactions. Therefore, specific identification of lectins (or bacteria) is complicated by the non-specific binding which can give false positives. To overcome this, we have reported the use of a powerful statistical analysis technique for the identification of carbohydrate

<sup>a</sup> Department of Chemistry, University of Warwick, Gibbet Hill Road, Coventry UK, CV4 7AL.

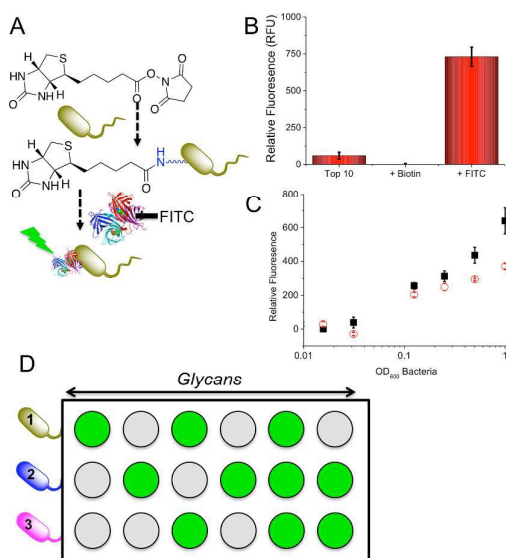
<sup>b</sup> School of Life Sciences, University of Warwick, Gibbet Hill Road, Coventry, UK.

\* Corresponding Author email; m.i.gibson@warwick.ac.uk

Electronic Supplementary Information (ESI) available: which includes experimental procedures. See DOI: 10.1039/x0xx00000x

binding toxins with similar binding specificity using linear discriminant analysis (LDA) – essentially we created a ‘bar code’ to described each lectin.<sup>32</sup> Here we demonstrate a multiplexed bacteria assay based on their relative binding affinity to a small and accessible array of glycans. The relative binding affinity of each bacteria was determined to generate a training matrix, which can subsequently be used for discrimination/identification. A range of Gram-negative, Gram-positive including mycobacteria are studied, including surrogates for *M. tuberculosis*.

The key aim here was to assess if the relative binding profiles of different bacterial strains to a range of glycans could be used to generate a training matrix as an easy method to enable multiplexed diagnostics based on each strains relative binding to different sugars. To generate glycosylated surfaces we employed hydrazide functional 96 well plates (Carbo-BIND™). Glycosylation with reducing carbohydrates was enabled by simply heating at 50 °C with an aniline catalyst. This method gives a mixture between ring-closed (approximately 60 %) and ring opened sugars. Whilst generating some heterogeneity, the discrimination method we employ (see later) is not affected by this, and in theory any other glycan immobilisation method could be used. To enable visualisation of bacterial binding for this proof-of-concept study, a two-step labelling procedure was employed. Bacteria were first biotinylated with Biotin-NHS, followed by reaction with FITC-labelled streptavidin (Figure 1A). We choose this method to enable facile binding measurement rather than as a true diagnostic tool. Alternative detection methods include microscopy, SPR, bilayer interferometry or nanoparticle binding, but are outside the scope of this manuscript. Successful labelling of the bacteria was showed by measuring the fluorescence (Figure 1B) of the bacterial cultures and visual examination.

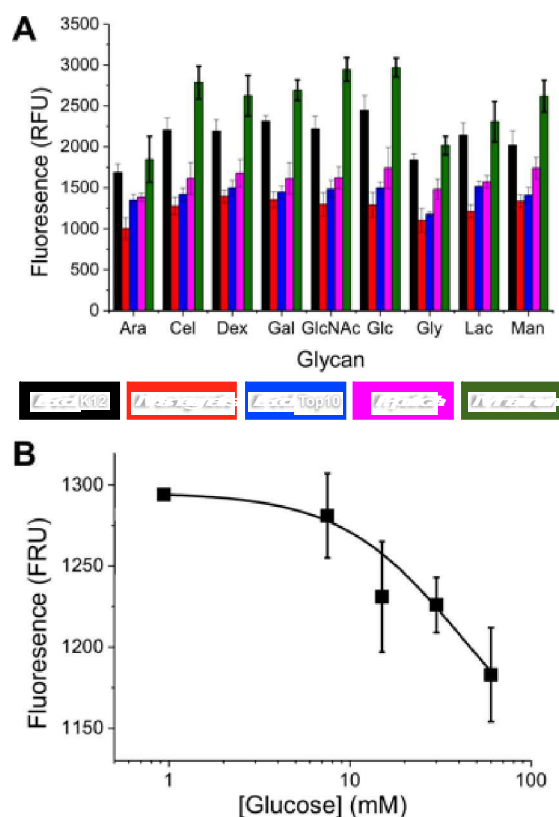


**Figure 1.** General concept employed here. A) Fluorescent labelling of bacteria; B) Demonstrating of dye conjugation; C) Binding to glycosylated surface of labelled Top10 (red circle) and K12 (black

square) *E. coli*; D) Discriminatory approach used here, where each bacteria will generate a different binding pattern.

To evaluate glycan binding, Top10 and K12 (*E. coli*) were incubated with a glucose functional surface as a function of bacterial concentration. A clear dose-dependant response was observed with the K12's showing more adherence at higher ODs, which we attribute to their expression of FimH (Figure 1C). This experiment also serves to highlight the challenge of single-channel glycan-based sensors – both bacteria would give a positive response. Figure 1D shows the multiplexed concept to be used here. Looking ‘across’ from any given glycan, it would not be possible to identify the bacteria. However, using multiple glycan ‘inputs’ enables a barcode to be created which is unique to the bacteria. It is therefore clear that several glycans are required to gain sufficient resolution. 9 different glycans were immobilised onto the plates and 5 different bacteria strains (labelled as described above) were interrogated. The bacterial species examined here are; two strains of *E. coli* Top10 and K12, *Pseudomonas putida*, *Mycobacterium smegmatis* and *Mycobacterium marinum*. Of the two strains of *E. coli*: Top10 is FimH negative and K12 is FimH positive and thus a mimic of pathogenic strains. *P. putida* is similar in nature to *Pseudomonas aeruginosa* which causes life threatening infections that can affect multiple organs but commonly causes pneumonia in cystic fibrosis patients.<sup>33, 34</sup> *M. marinum* and *M. smegmatis* are both model organisms for *M. tuberculosis*, the causative agent of tuberculosis, with *M. marinum* sharing around 3000 orthologs with over 85 % amino acid identity with *M. tuberculosis*.<sup>35</sup> The species analysed included Gram-positive and Gram-negative species, pathogenic and non-pathogenic samples and different strains of the same species providing a focussed, but diverse set of bacteria. The relative binding of these bacteria to each surface was evaluated by fluorescence and the profiles are show in Figure 2A.

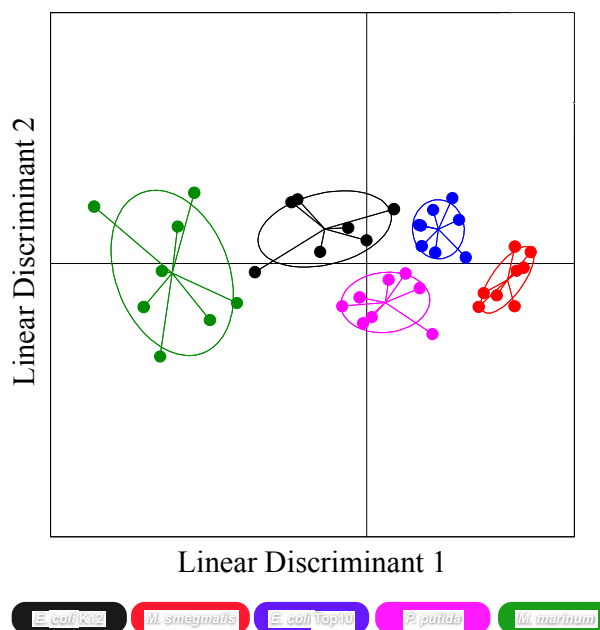
Figure 2A shows that individual strains exhibited distinct levels of adhesion. For example *M. marinum* showed higher binding to most glycans than the others. It is beyond the scope of this study to identify the individual components responsible for the binding, and it should be highlighted that this knowledge is not essential for this biosensory approach (*vide infra*). However, it was important to confirm the interactions were due to glycan-adhesion, and not just non-specific binding. A competition experiment was conducted whereby OD = 1 *M. smegmatis* was incubated with the surfaces, but with increasing glucose concentrations in the buffer (1 – 50 mM). As the glucose concentration was increased, there was a clear decrease in the amount of binding (Figure 2B).



**Figure 2.** A) Binding profiles of fluorescently labelled bacteria to glycan surfaces. Arabinose (Ara), cellobiose (Cel), dextran (Dex), galactose (Gal), glucose (Glc), N-Acetyl-D-glucosamine (GlcNAc), glyceraldehyde (Gly), lactose (Lac) and mannose (Man). Each bar represents the average of 8 replicates and the error bars represent the standard error; B) Competitive binding of *M. smegmatis* to glucose surface, in presence of free glucose. Error is S.D from minimum of 3 repeats.

There are many putative carbohydrate binding proteins (including transporters) in the strains, and any given adhesin can also bind to other glycans, even if it has relatively lower affinity. Considering this, any single sugar cannot identify a specific bacteria, within the error of the measurement. However, each bacteria has a distinct set ('barcode') of binding across the range of sugars. This large set of data was thus used as a training matrix for LDA.<sup>32, 36</sup> LDA is a powerful classification tool, which for every class (bacterial species) within the training matrix (binding profiles) it minimises the variation within each class and maximises the variation between classes by transforming the original data in order to get maximum separation between groups. For the bacterial binding data, the LDA model produced showed good resolution between bacterial species and when validated using a 'leave-one-out' approach (where each sample is left out in an iterative process and the model generated before classification of the missing sample) the model was able to re-classify all data points to their original class with 92 % accuracy. Figure 3 shows a 2-dimensional representation of the LDA analysis and confidence boundaries for each bacterial strain. This clearly

shows that each bacteria has a distinct 'region' enabling discrimination. In particular, the two mycobacterial species (*M. smegmatis* and *M. marinum*) had excellent separation.



**Figure 3.** Linear discriminant analysis model of bacterial binding. *E. coli* K12, *E. coli* Top10, *M. marinum*, *P. putida* and *M. smegmatis*. Each point represents a single binding profile to all the glycan surfaces, the centre point dictates the average value for all samples and the ellipses represent one standard deviation from the average.

Finally a blind test was done. A culture was prepared by an independent operator and provided. This was fluorescently labelled and incubated with each of the glycan surfaces in triplicate. The average binding response was classified using the LDA model and was correctly identified as *E. coli* Top10 with 96 % certainty. Future work will involve the extension of this method to 'label-free' analytical techniques and also focussing on clinical strains. The rapid and simple nature of this makes it ideal for translation and compares favourable with sequencing/PCR based methods such as Ibis which takes at least 6 hours for identification<sup>14</sup>, and in some cases expansion and culture of the bacteria.

To conclude, we have demonstrated a rapid technique for the identification of bacterial strains based on profiling their differential binding to carbohydrate functionalised surfaces coupled with a powerful, but simple to use, classification algorithm. The procedure was shown to be valid for Gram-negative, Gram-positive and mycobacteria, indicating it has potential to diagnose or identify a broad range of pathogenic bacteria, either on its own, or in combination with other established analytical methods. In particular, the potential for rapid and low-cost diagnostics is shown including for mycobacteria. *M. tuberculosis* is a re-emerging global healthcare threat with no suitable detection methods.

## Notes and references

We would like to thank Dr Alasdair Hubbard for technical assistance. Equipment was supported by the Birmingham Science City (SC) Advanced Materials project, with support from Advantage West Midlands and part funded by the European Regional Development Fund. LO is funded by the BBSRC via the Systems Biology DTC. EF is a Sir Henry Dale Fellowship jointly funded by the Wellcome Trust and the Royal Society (Grant Number 104193/Z/14/Z). MIG thanks the ERC for a Starting Grant, CRYOMAT 638661.

1. M. T. Holden, L. Y. Hsu, K. Kurt, L. A. Weinert, A. E. Mather, S. R. Harris, B. Strommenger, F. Layer, W. Witte, H. de Lencastre, R. Skov, H. Westh, H. Zemlickova, G. Coombs, A. M. Kearns, R. L. Hill, J. Edgeworth, I. Gould, V. Gant, J. Cooke, G. F. Edwards, P. R. McAdam, K. E. Templeton, A. McCann, Z. Zhou, S. Castillo-Ramirez, E. J. Feil, L. O. Hudson, M. C. Enright, F. Balloux, D. M. Aanensen, B. G. Spratt, J. R. Fitzgerald, J. Parkhill, M. Achtman, S. D. Bentley and U. Nubel, *Genome Research*, 2013, **23**, 653-664.
2. C. M. Surawicz, L. J. Brandt, D. G. Binion, A. N. Ananthakrishnan, S. R. Curry, P. H. Gilligan, L. V. McFarland, M. Mellow and B. S. Zuckerbraun, *American Journal of Gastroenterology*, 2013, **108**, 478-498.
3. A. Zaidi, W. Huskins, D. Thaver, Z. Bhutta, Z. Abbas and D. Goldmann, *The Lancet*, 2005, **365**, 1175-1188.
4. J. M. Q. Hiram, F. Ricardo, C. Miguel, A. Jose, C. B. Maria, C. M. S. Juan, R. Pedro Torellas and C. Veronica, in *B48. case reports on Tuberculosis*, American Thoracic Society, 2015, pp. A3275-A3275.
5. N. Rezaei, A. Aghamohammadi, D. Mansouri, N. Parvaneh and J. L. Casanova, *Expert Review on Clinical Immunology*, 2011, **7**, 129-131.
6. C. M. Grossman, *Annals of Internal Medicine*, 2008, **149**, 135-136.
7. M. Cully, *Nature*, 2014, **509**, S16-S17.
8. D. Bethune, R. Blowers, M. Parker, E. Pask, O. Jessen, K. Rosendal, P. Bulow, V. Faber, K. Eriksen and K. Hiramatsu, *British Medical Journal*, 1961, **124**, 124-125.
9. J. O'Neil, Antimicrobial resistance: Tackling a crisis for the health and wealth of nations., <http://amr-review.org/home>.
10. B. Spellberg, R. Gidos, D. Gilbert, J. Bradley, H. W. Boucher, W. M. Scheld, J. G. Bartlett, J. Edwards and T. I. D. S. o. America, *Clinical Infectious Diseases*, 2008, **46**, 155-164.
11. P. Tissari, A. Zumla, E. Tarkka, S. Mero, L. Savolainen, M. Vaara, A. Aittakorpi, S. Laakso, M. Lindfors, H. Piiparinen, M. Mäki, C. Carder, J. Huggett and V. Gant, *The Lancet*, **375**, 224-230.
12. S. Tan, S. Chew, S. Tan, M. Givskov and L. Yang, *Current Opinion in Biotechnology*, 2014, **26**, 1-6.
13. R. Koch and T. Brock, *Milestones in microbiology: 1556 to 1940*, 1884.
14. J. W. Costerton, J. C. Post, G. D. Ehrlich, F. Z. Hu, R. Kreft, L. Nistico, S. Kathju, P. Stoodley, L. Hall-Stoodley, G. Maale, G. James, N. Sotereanos and P. DeMeo, *FEMS Immunology & Medical Microbiology*, 2011, **61**, 133-140.
15. S. Pasquaroli, G. Zandri, C. Vignaroli, C. Vuotto, G. Donelli and F. Biavasco, *Journal of Antimicrobial Chemotherapy*, 2013, **68**, 1812-1817.
16. E. L. Doughty, M. J. Sergeant, I. Adetifa, M. Antonio and M. J. Pallen, *PeerJ*, 2014, **2**, e585.
17. M. R. Halachev, J. Z. Chan, C. I. Constantinidou, N. Cumley, C. Bradley, M. Smith-Banks, B. Oppenheim and M. J. Pallen, *Genome Medicine*, 2014, **6**, 70.
18. M. Hartmann, H. Papavlassopoulos, V. Chandrasekaran, C. Grabosch, F. Beiroth, T. K. Lindhorst and C. Röhl, *FEBS letters*, 2012, **586**, 1459-1465.
19. S. Sakarya and S. Oncu, *Medical Science Monitor*, 2003, **9**, RA76-82.
20. N. P. Pera and R. J. Pieters, *Medicinal Chemistry Communication*, 2014, **5**, 1027-1035.
21. A. K. Ojha, A. D. Baughn, D. Sambandan, T. Hsu, X. Trivelli, Y. Guerardel, A. Alahari, L. Kremer, W. R. Jacobs and G. F. Hatfull, *Molecular Microbiology*, 2008, **69**, 164-174.
22. L.-C. Xu and C. A. Siedlecki, *Acta biomaterialia*, 2012, **8**, 72-81.
23. J. W. Costerton, P. S. Stewart and E. P. Greenberg, *Science*, 1999, **284**, 1318-1322.
24. B. Hoyle and J. W. Costerton, in *Progress in Drug Research*, ed. E. Jucker, Birkhäuser Basel, 1991, vol. 37, ch. 2, pp. 91-105.
25. M. Jones, L. Otten, S.-J. Richards, R. Lowery, D. Phillips, D. Haddleton and M. Gibson, *Chemical Science*, 2014, **5**, 1611-1616.
26. S. J. Richards, M. W. Jones, M. Hunaban, D. M. Haddleton and M. I. Gibson, *Angewandte Chemie International Edition*, 2012, **51**, 7812-7816.
27. S.-J. Richards, E. Fullam, G. S. Besra and M. I. Gibson, *Journal of Materials Chemistry B*, 2014, **2**, 1490-1498.
28. R. L. Phillips, O. R. Miranda, C. C. You, V. M. Rotello and U. H. Bunz, *Angewandte Chemie International Edition*, 2008, **47**, 2590-2594.
29. M. D. Disney and P. H. Seeberger, *Chemical Biology*, 2004, **11**, 1701-1707.
30. E. Bulard, A. Bouchet-Spinelli, P. Chaud, A. Roget, R. Calemczuk, S. Fort and T. Livache, *Analytical Chemistry*, 2015, **87**, 1804-1811.
31. N. P. Pera, A. Kouki, S. Haataja, H. M. Branderhorst, R. M. J. Liskamp, G. M. Visser, J. Finne and R. J. Pieters, *Organic & Biomolecular Chemistry*, 2010, **8**, 2425-2429.
32. L. Otten and M. I. Gibson, *RSC Advances*, 2015, **5**, 53911-53914.
33. A. Wittgens, T. Tiso, T. T. Arndt, P. Wenk, J. Hemmerich, C. Muller, R. Wichmann, B. Kupper, M. Zwick, S. Wilhelm, R. Hausmann, C. Sylđatk, F. Rosenau and L. M. Blank, *Microbial Cell Factories*, 2011, **10**, 80.
34. N. Hoiby, E. W. Flensburg, B. Beck, B. Friis, S. V. Jacobsen and L. Jacobsen, *Scandinavian Journal of Respiratory Disease*, 1977, **58**, 65-79.
35. T. P. Stinear, T. Seemann, P. F. Harrison, G. A. Jenkin, J. K. Davies, P. D. Johnson, Z. Abdellah, C. Arrowsmith, T. Chillingworth, C. Churcher, K. Clarke, A. Cronin, P. Davis, I. Goodhead, N. Holroyd, K. Jagels, A. Lord, S. Moule, K. Mungall, H. Norbertczak, M. A. Quail, E. Rabinowitsch, D. Walker, B. White, S. Whitehead, P. L. Small, R. Brosch, L. Ramakrishnan, M. A. Fischbach, J. Parkhill and S. T. Cole, *Genome Research*, 2008, **18**, 729-741.

Journal Name

COMMUNICATION

36. A. Izenman, *Modern multivariate statistical techniques*, Springer, 2008.