Supporting Information

Generation and Characterization of a Library of Novel Biologically Active Functional Surfactants (Surfmers) using Combined High-Throughput Methods

Valentina Cuzzucoli Crucitti[a], Leonardo Contreas[b], Vincenzo Taresco[c], Shaun C. Howard[d], Adam A. Dundas[a], Marion J. Limo[e], Takasi Nisisako[f], Philip M. Williams[b], Paul Williams[g], Morgan R. Alexander[b], Ricky D. Wildman[a], Benjamin W. Muir[d]* and Derek J. Irvine[a]**

[a] Centre for Additive Manufacturing and Department of Chemical and Environmental Engineering, Faculty of Engineering, University of Nottingham, NG7 2RD, Nottingham, UK

Email:

[b] School of Pharmacy, University of Nottingham, NG7 2RD, Nottingham, UK

[c] University of Nottingham, NG7 2RD, Nottingham, UK

[d] CSIRO Manufacturing, Clayton, 3168, Victoria, Australia

[e] Interface and Surface Analysis Centre, University of Nottingham, NG7 2RD, Nottingham, UK

[f] Institute of Innovative Research, Tokyo Institute of Technology, Yokohama, Kanagawa, 226-8503, Japan

[g] Biodiscovery Institute, National Biofilms Innovation Centre and School of Life Sciences, University of Nottingham, NG7 2RD, UK

*Ricky D. Wildman Email: Ricky.Wildman@nottingham.ac.uk

*Benjamin W. Muir Email: Ben.Muir@csiro.au

*Derek J. Irvine Email: Derek.Irvine@nottingham.ac.uk

S1 Experimental Section

Materials: All the materials were used as received unless stated otherwise. The monomeric species used to build the library of hydrophobic monomers were purchased from the following companies Sigma-Aldrich, Combi-blocks and Polysciences (Table S1). Poly(ethylene glycol) methyl ether methacrylate (mPEGMA) (Mn = 300) and 2, 2′-Azobis (2-methylpropionitrile) (AIBN, 98%) and the Benzyl Mercaptan (BM, 99%) were also was purchased from Sigma Aldrich. The catalytic chain transfer agent, Bis[(difluoroboryl)diphenylglyoximato] cobalt (II) (PhCoBF), was supplied from DuPont. The cyclohexanone and heptane used as solvents for the synthesis and precipitations respectively and were used as received from and supplied by Fisher Scientific.

Table S1. List of vendors, purity and acronyms of the monomers used.

Entry	Monomers	Acronyms	Supplier	Purity
1	Ethylene glycol dicyclopentenyl ether acrylate	EGDPEA	S-Aldrich	≥90%
2	Isobornyl methacrylate	iBMA	S-Aldrich	≥80%
3	Phenyl acrylate	PhA	Polysciences	95%
4	2-N-Morpholinoethyl methacrylate	NMEMA	S-Aldrich	95%
5	Tetrahydrofurfuryl acrylate	THFuA	S-Aldrich	≥90%
6	2-(Methacryloyloxy)ethyl acetoacetate	MAEA	S-Aldrich	95%
7	n-Butyl Acrylate	BuA	S-Aldrich	≥99%
8	Ethyl Acrylate	EA	S-Aldrich	99%
9	Laury acetate	LaA	S-Aldrich	≥90%
10	N-[3(Dimethylamino)propyl]methacrylamide	DMPAm	S-Aldrich	99%
11	Heptafluorobutyl methacrylate	F7BMA	Combi-block	95%
12	n-Hexyl acrylate	HA	S-Aldrich	98%
13	Hydroxy-3-phenoxypropyl acrylate	HPhOPA	S-Aldrich	≥80%
14	Ethylene glycol phenyl ether acrylate	EGPhEA	S-Aldrich	≥82%
15	3-(Dimethylamino)propyl acrylate	DMAPA	S-Aldrich	95%
16	Isobutyl methacrylate	iBuMA	S-Aldrich	97%
17	Furfuryl methacrylate	FuMA	S-Aldrich	97%
18	n-Decyl methacrylate	DMA	Polysciences	99%
19	Phenyl methacrylate	PhMA	S-Aldrich	90%
20	Isobutyl acrylate	iBuA	S-Aldrich	≥99%

High-Throughput Reactors Configuration: A Chemspeed Swing robot equipped with an isynth reactor containing 48 individual reactors was used for all polymerizations. The isynth reactor was fitted with 8 mL disposable glass vials obtained from Chemspeed Technologies Pty Ltd. The enclosed Chemspeed robotic deck was made inert via constant Nitrogen flow over 50

L/min to remove all oxygen with the extraction ports closed. Nitrogen purged reagent vials were passed into the robot immediately after the completion of degassing using a transfer chamber with a vacuum set to 200 mbar to minimize any exposure to oxygen. All reagent vials were degassed with nitrogen for a minimum of one hour prior to transferring into the deck of the robot.

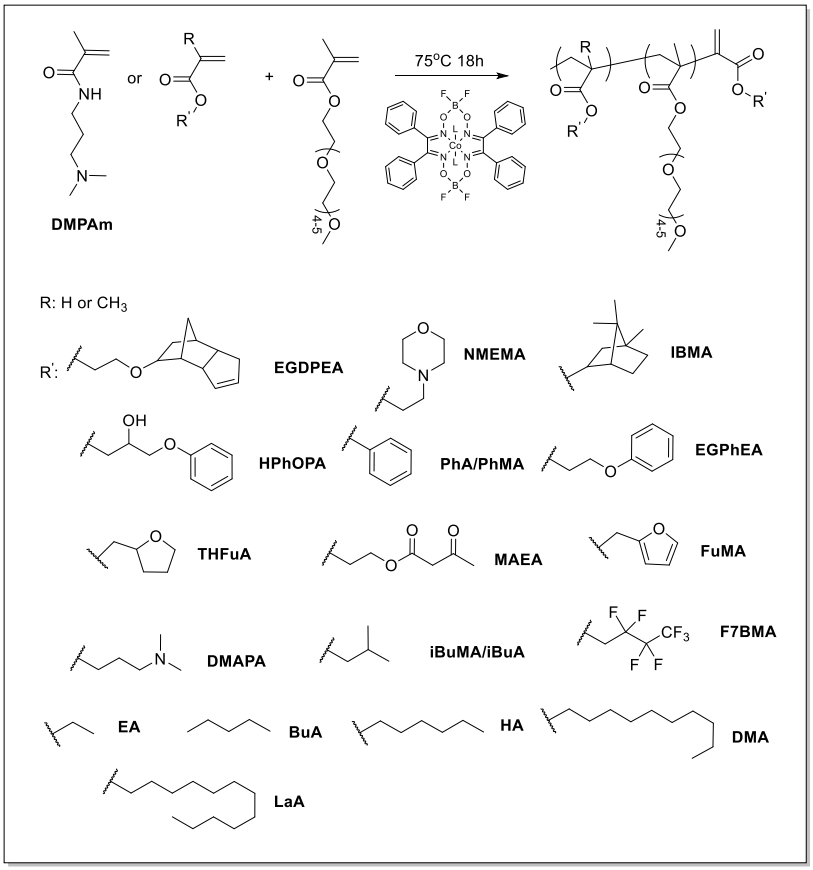
For all aspirations and dispensing of reagent solutions, a 4-Needle Head tool equipped with 2 x 1 mL and 2 x 10 mL syringes was used which was fitted with stainless steel septa piercing needles. All solvent lines were primed with degassed cyclohexanone which was used for each rinsing step. Typical aspiration and dispense rates of the reagents were 2 ml/min and 5 ml/min, respectively, for the 1 mL syringes and 10 ml/min and 5 ml/min, respectively, for the 10 mL syringes. An airgap of 50 µL and an extra volume of 50 µL were used for all aspirations using the 4-Needle Head tool. The needles were rinsed after each reagent dispense step with a 3 mL inside and outside volume for the 1 mL syringes and a 5 mL inside and outside volume for the 10 mL syringes. All reagents were added to the reactors including monomers, CTA, initiator and solvent prior to heating the isynth reactor under reflux conditions to obtain a temperature of 75 ℃ inside the reactors for 18 hours. The isynth lid was set to closed independent for the polymerisations while being actively cooled to 4 degrees Celsius under reflux. The isynth reactor was set to shake at 400 rpm for the duration of the polymerizations to ensure adequate mixing. It was then cooled to 20 ℃ in order to cease the polymerizations.

The scale up of a small range of copolymers was performed on a Chemspeed SLT2 robot equipped with an isynth reactor block containing 12 individual reactors. The isynth reactor was fitted with 100 mL disposable glass vials obtained from Chemspeed Technologies Pty Ltd. A 250 ml reagent bottle of cyclohexanone was placed into the robot with continuous nitrogen bubbling throughout the experiment (minimum 2 hours before use). Septa-capped, nitrogen purged reagent

vials were placed into the robot immediately after the completion of degassing (minimum 1 hour), and the enclosed Chemspeed robotic deck was made inert via constant Nitrogen flow over 50 L/min to remove all oxygen with the extraction ports closed. The internal atmosphere was reduced to 0.1% oxygen within 90 minutes. For all aspirations and dispensing of reagent solutions, a 4-Needle Head tool equipped with 2 x 1 mL and 2 x 10 mL syringes was used which was fitted with stainless steel septa piercing needles. All solvent lines were primed with degassed cyclohexanone, which was used for each rinsing step. Typical aspiration and dispense rates of the reagents were 20 and 40 ml/min (respectively) for the 10 mL syringes (the 1 mL needles were not used for dispensing). An airgap of 50 uL and an extra volume of 50 uL was used for all aspirations using the 4-Needle Head tool. The needles were rinsed after each reagent dispense step with a 20 mL inside and outside volumes. Prior to heating the Isynth, reactors were agitated for 4 minutes, and a 250 uL sample was taken from each reactor and automatically dispensed into NMR tubes for later analysis. This was accomplished using 2 x 1 mL needles on the 4-Needle head tool, with 2 mL inside and outside volume rinse following each sample.

The purifications of the polymers were conducted in excess of either heptane or methanol according to the low solubility of the materials in such solvents. The usual non-solvent:reaction media ratio was 5:1 vol/vol in order to enhance the precipitation process and, finally, the precipitated materials were collected in a vial and left in vacuum oven for at least 24hrs.

General Procedure Adopted for the Initial Parallel Polymerization Screen Contemplating 54 Polymerizations using CCTP at 1.5 g scale:



Scheme S1. CCTP reaction mechanism and the different monomers employed in the HT screening.

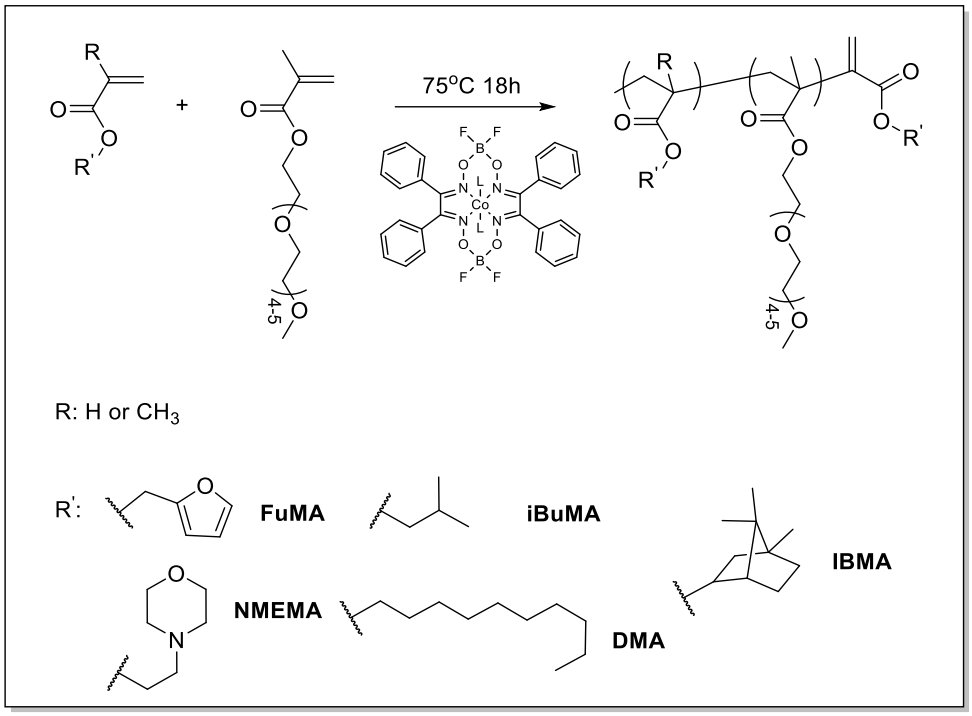
The general procedure adopted for the polymerizations was as follows. Stock Solutions of the hydrophobic monomers were prepared covering a range between 3 M, 3.5 M and 4.5 M. The concentrations of stock solutions of $mPEGMA_{300}$, initiator and catalyst (with respect to the monomers) were 1.5 M, 0.1 M and 3 mg/ mL, respectively.

The hydrophobic monomer and $mPEGMA_{300}$ were automatically dispensed into each reactor in the appropriate quantities required to reach the targeted molar ratio 90:10 hydrophobic monomer/$mPEGMA_{300}$ as defined by the pre-programming of the Chemspeed. The concentrations of the catalyst and initiators adopted for the first screening were fixed for all the co-polymerizations at 850 ppm and 0.5% w/w, respectively. The only parameter that varied in this screen was the solvent ratio and the concentration of the reaction media. The solvent:monomer

ratios considered for this study were 1:3 v/v and 1:2 v/v (monomers:cyclohexanone) for 17 monomers, while, in the case of the 2-Hydroxy-3-phenoxypropyl acrylate (HPhOPA) and Ethylene glycol phenyl ether acrylate (EGPhEA) lower concentrations were also explored using the ratios 1:4 v/v and 1:5 v/v (monomers:cyclohexanone).

This sequence was carried out in duplicate in order to confirm the reproducibility of the polymerisation using CCTP for conversion and the control of the molecular weight.

Optimization of 10 CCTP Parallel Polymerization which exhibited no conversion in the initial screening at 1.5g scale:

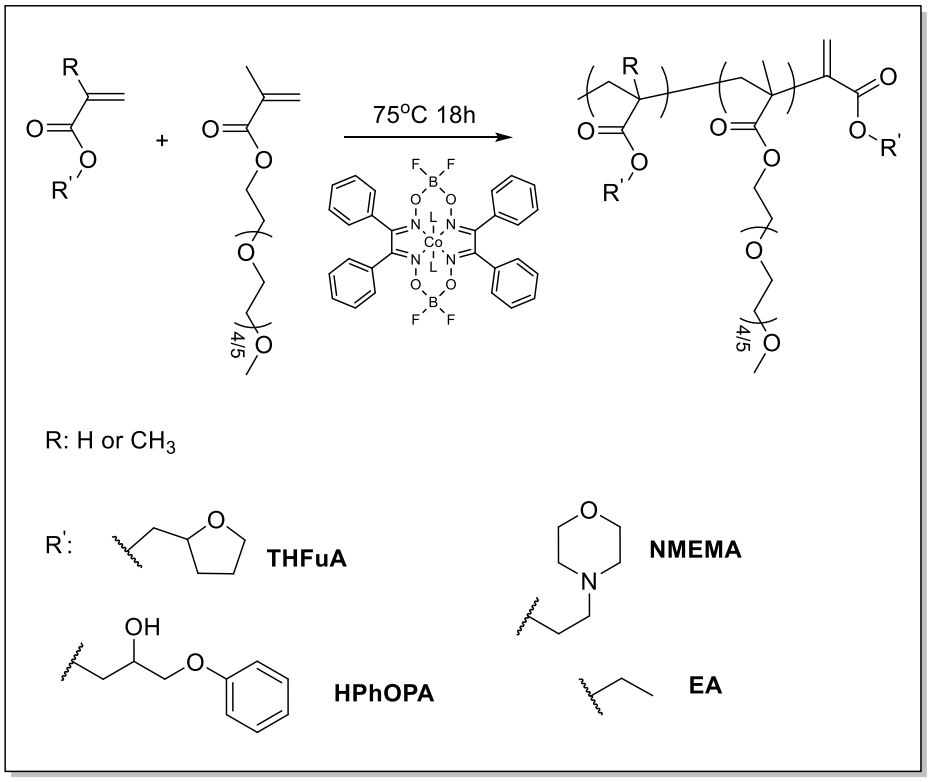


Scheme S2. CCTP reaction mechanism and the different monomers employed in the HT screening optimization.

This second attempt at using the CCTP in automated fashion was related to the optimization of the catalyst concentrations for those monomers that showed either no or low conversion (< 30%) in the first sequence. The new concentrations of PhCoBF explored were 650 ppm and 450 ppm

relative to the monomers. The same procedures as detailed for the 44 parallel polymerizations was utilized. The feed ratio of hydrophobic monomer to $mPEGMA_{300}$ was 90:10 mol/mol (i.e. HB:$mPEGMA_{300}$). Stock solutions of the hydrophobic monomer spanned a range from 3 M to 4.5 M, while, for the $mPEGMA_{300}$, AIBN and PhCoBF the concentrations used were 2 M, 0.03 M and 2 mg/mL, respectively. Based on the information gathered from the first screening, the solvent ratio preferred was the 1:3 vol/vol (monomer:cyclohexanone) and the initiator was employed at 0.5% wt/wt.

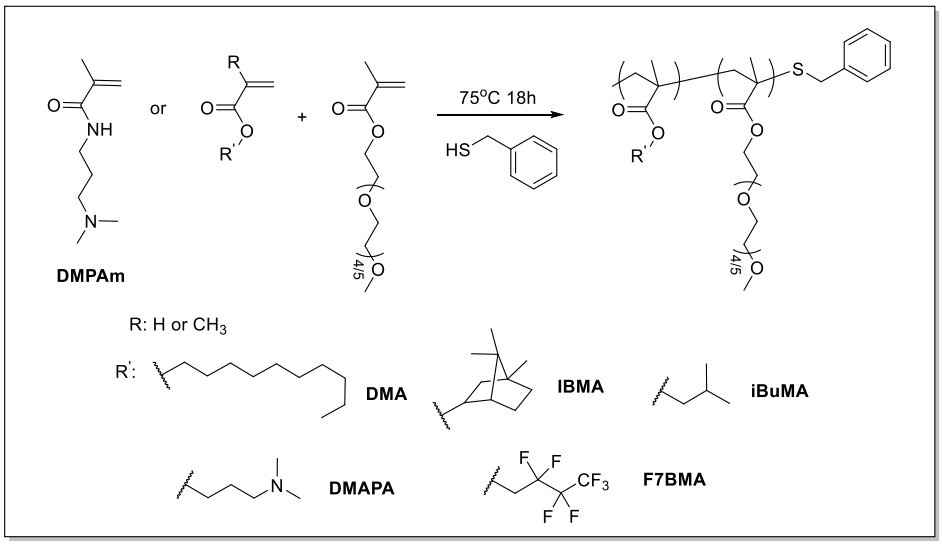
Scale up of 4 selected Parallel Polymerization conducted via CCTP from 1.5g to 10 g scale:



Scheme S3. CCTP reaction mechanism and the different monomers employed in the HT scale-up.

Four copolymers were chosen from the libraries obtained during the initial screening and synthesis optimization stages to be up scaled to a product volume of 10 g via the Chemspeed robotic synthesizer. These reactions were conducted using those conditions found to promote both the highest polymer conversion and deliver molecular weights (MWt) close to the chosen target. The conditions (i.e. solvent ratios and catalyst ppm) applied were defined from the results of the screening polymerizations. As in the previous experiments, the different stock solutions were pre-produced containing hydrophobic monomer and $mPEGMA_{300}$, initiator and catalyst. In this case, an initiator concentration of 0.5% wt/wt and HbM:$mPEGMA_{300}$ ratio of 90:10 mol/mol were adopted and fixed for the 4 polymerizations.

General procedure adopted for the Parallel Polymerization of 12 polymerizations conducted using a Thiol Control Agents at 1.5 g scale:



Scheme S4. Thiol-mediated Free Radical Polymerization reaction mechanism and the different monomers employed in the HT scale-up.

The same basic procedure steps were applied for the thiol-controlled polymerizations as detailed for the CCTP polymerisations. Specifically, stock solutions of all the reagents were prepared as described above, to the same concentrations with the exception of that of the CTA. The thiol species chosen as CTA for these experiments was benzyl mercaptan (BzSH) of which two concentrations were applied in the attempt of the control of molecular weight, 10% mol/mol and 5% mol/mol (with respect to the total monomer concentrations). The solvent ratio applied was 1:3 v/v, the AIBN was introduced at the concentration of 0.5% wt/wt and the monomers ratio was 90:10 mol/mol.

Characterization of Products from the High-Throughput Polymer Libraries: *Standard Throughput ^{1}H NMR Analysis.* ^{1}H NMR spectra were recorded at 25°C using a Bruker DPX-300 spectrometer (400 MHz) and Chemical shifts were recorded in δH (ppm). Samples were dissolved in deuterated chloroform ($CDCl_3$) to which chemical shifts are referenced (residual chloroform at 7.26 ppm). The ^{1}H-NMR analysis was performed on the crude polymerization solutions to determine conversion. By using the Chemspeed robot, 250 uL aliquots were sampled at 0 and 18 hours each reactor using the 4-Needle Head tool from, which was then dispensed directly into the NMR tube, and, d- chloroform was then added. Conversions were calculated by considering the decrease in the integral values related to the double bond characteristic of (meth)acrylate monomers in the ^{1}H-NMR spectra from the start to the end points. In order to have a reference in both the NMR spectra (0 and 18 hrs) tetramethyl silane (TMS) was added to the d-chloroform in a known amount. The product polymers were subsequently purified following the method defined and the final composition of the copolymers was defined from the ^{1}H-NMR spectrum. The formula applied for the determination of the percentage conversion (i.e % MON) was that detailed as **Equation S1**:

$$\%\mathrm{MON} = \frac{\frac{a_x}{m_x}}{\frac{a_x}{m_x}+\frac{a_y}{m_y}} * 100 \qquad \text{(Equation S1)}$$

Equation S1. Equation used to determine the final hydrophobic monomer:mPEGMA$_{300}$ final ratio. a_x and a_y are the values of the peaks area related to both the monomers, while, m_x and m_y are the number of protons that are related to each of these resonances.

The OCH_3 resonance of the mPEGMA$_{300}$ macromonomer, that occurs at δ 3.4 ppm, was used in the calculation of the actual monomer ratio as it remains unaltered in the hydrophilic component for all the copolymers produced (**Figure S1**).

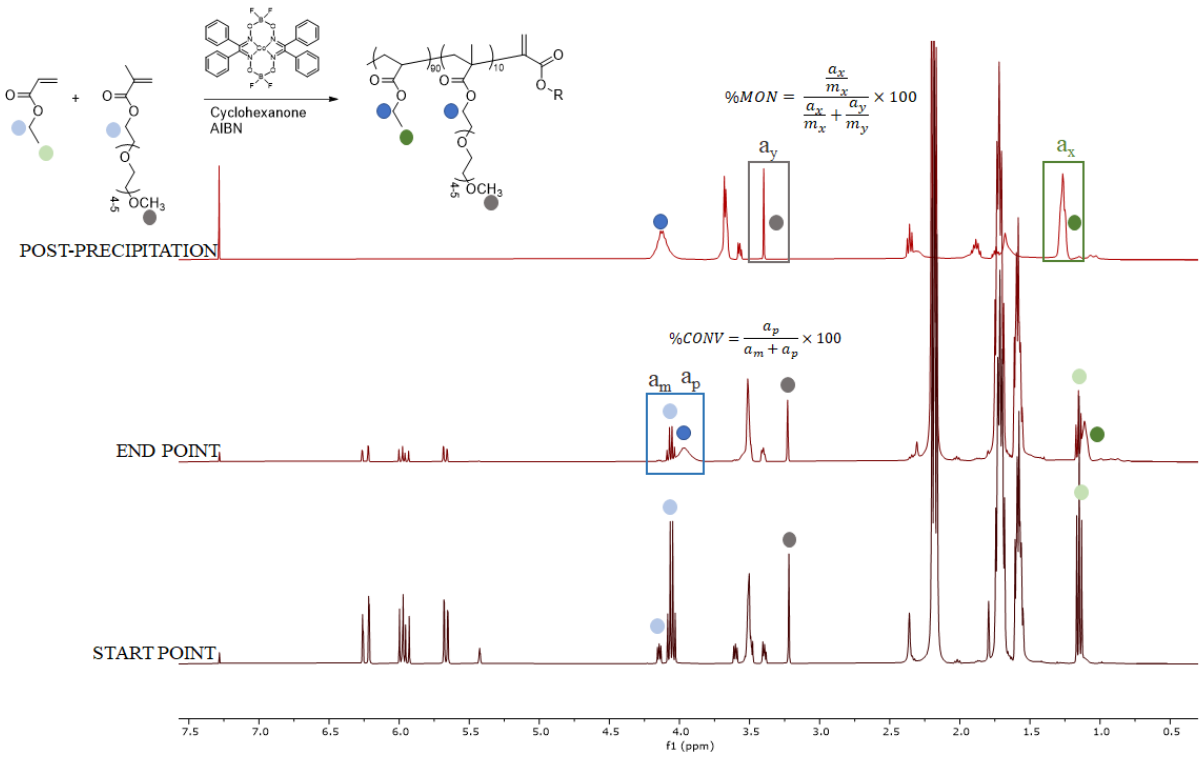


Figure S1. Example to show how the final HbM:mPEGMA$_{300}$ molar ratio was calculated from the ^{1}H-NMR. The copolymer used for this example is EA-co-mPEGMA$_{300}$.

Standard Throughput Gel Permeation Chromatography (GPC) Analysis: Identical volumes to that used for the NMR analysis was also withdrawn by the robot from each reactor and directly dispensed into GPC vials for analysis to which HPLC grade THF was then added to dissolve the sample. GPC was performed on a Waters Alliance system equipped with an Alliance 2695 Separations Module (integrated quaternary solvent delivery, solvent degasser and autosampler system), a Waters column heater module, a Waters 2414 RDI refractive index detector, a Waters PDA 2996 photodiode array detector (210 to 400nm at 1.2nm) and 4× Agilent PL-Gel columns (3 x PL-Gel Mixed C (5um) and 1 x PL-Gel Mixed E (3um) columns), each 300 mm × 7.8 mm2, providing an effective molar mass range of 200 to 2 × 10^6). Tetrahydrofuran (THF) high purity solvent (HPLC grade) was pre-filtered through aluminium oxide (90 active neutral, 70-230 mesh) with 0.45um filter, and 0.1 g L-1 2,6-di-tert-butyl-4-methylphenol(BHT) was added as inhibitor. The filtered THF containing BHT was purged slowly with nitrogen gas and used as an eluent with a flow rate of 1 mL/min at 30°C. Number (Mn) and weight average (Mw) molar masses were evaluated using Waters Empower-3 software. The GPC columns were calibrated with low dispersity polystyrene (PSt) standards (Polymer Laboratories) ranging from 580 to 7,500,000 g mol-1, and molar masses are reported as PSt equivalents.

High-Throughput Critical Micellar Concentration Analysis (CAC): The CAC HT analysis strategy combined three procedures in series; (a) preparation of stock nanoaggregate emulsions (NA) by nanoprecipitation, (b) dispersion of these NAs at different concentrations into a 96-well plate by means of an ink-jet printer and (c) measurement of the scattered light intensity (Cnt/s) with an automated DLS plate reader. The initial surfactant emulsions were prepared by a nanoprecipitation method using THF as the solvent and water as the non-solvent. The THF polymeric solutions (5 mg in 2 mL) were added manually *via* a syringe dropwise into milliQ water

(10 mL) in order to obtain a final emulsion concentration of 500 µg/mL. The setup of the dispensing process and target area definition are summarised in the Figure S2.

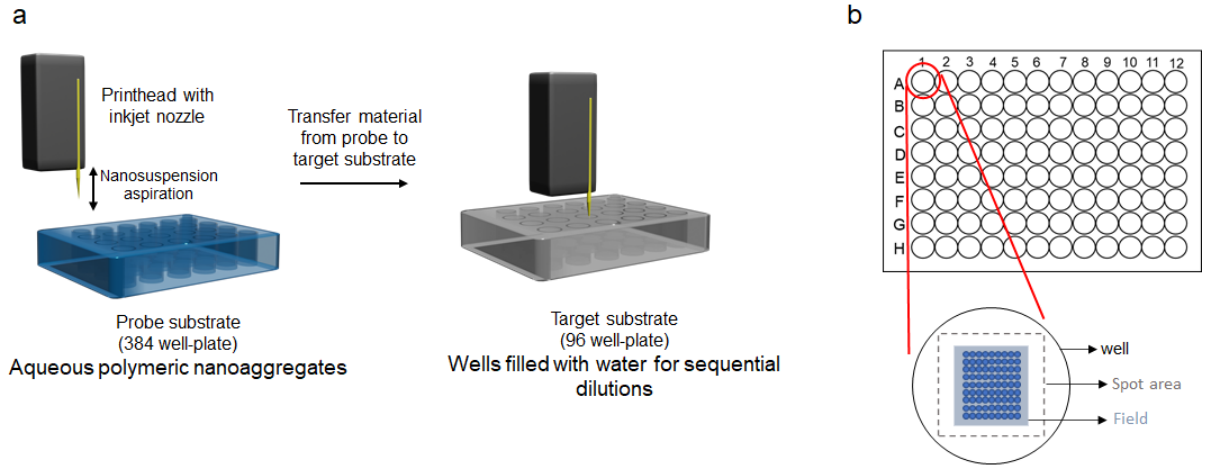


Figure S2. (a) Printing process by focusing on the nozzle setup. On the left-hand side, the aqueous polymeric nanosuspension/aggregation aspiration. On the right-hand side, printing dispensing of the aspired nanoaggregates into wells filled with 200 µL of water. (b) the 96-well plate target substrate, detailed view of a well with designed spot area and field, selected print pattern with adjustable number of drops per spot in magnified view.

Figure S2 shows that prior to dispensing the aqueous polymeric suspensions/emulsions into a 96-well plate filled with water, (200 µL per well), the target was defined. The plate dimensions were inserted into the sciFLEXARRAYER software (Scienion AG, version 2.09.002), subsequently the number of wells, well distance, well depth, and the spot area (area within the target designated for spotting) were also defined. Within the spot area, a field pattern was defined as a 9 × 9 spot matrix. After setting up the field pattern, the field setup was used to address the wells from the probe substrate to the spots. The probe substrate consists of the well plate containing the polymeric suspensions/emulsions with a starting concentration of 1 mg mL^{-1}. The number of

drops delivered per well were selected in such way that the volume aspirated for delivery by the nozzle (max 10 μL) at the beginning of a run was sufficient to print the whole print pattern.

The Aqueous polymeric suspensions/emulsions were dispensed by a piezoelectric inkjet printer (Sciflexarray S5, Scienion) using a 90 μm orifice nozzle. The nozzle was programmed to dispense the polymeric suspensions/emulsions into the well from a vertical distance of ≈10–20 mm from the well plate, avoiding touching the water surface in the target well. In all experiments, a defined field (Figure S2b) was selected as the "print pattern." The droplet size was controlled by the values of the voltage and electrical pulse. The voltage and electrical pulse were also tuned to prevent the occurrence of satellite. In a routine experiment, water suspension (500 μg/mL) droplets with nominal volumes ranging from 180-220 pL, were dispensed at a 300 Hz jetting frequency by adjusting the voltage and pulse between 98-105 Volt (Voltage) and 45-55 μs (Pulse) respectively. The nozzle was washed with water, in between each printing cycle, as part of the automated printing-washing loop. In between each printing cycle, the nozzle was washed with DMF as part of the automated printing/washing loop.

Images of the drop formation and droplet size were obtained using the printer software. The final spots were imaged using the Leica MZ16 stereomicroscope. The number of drops per spot was selected in such a way that the volume aspired by the nozzle (max 10 μL) at the beginning of each run was sufficient to dispense the whole (or multiple) targeted final dilutions of the nanoaggregates in target well.

The intensity (Cnt/s) and size (nm) of the light scattered by the final printed nanosuspensions was measured using a Wyatt DynaPro Plate Reader II DLS instrument, which has a laser wavelength of 817.28 nm and a scattering angle of 158°. The experimental procedures for the intensity measurements were conducted with the temperature fixed at 25 °C. Additionally, a fixed

laser power intensity and attenuation were applied in order to determine the scattered light from each sample. By comparison, in the size assessment experiments that were conducted on the same well-plate, auto-attenuation was enabled to determine the optimal laser power and attenuation. In these CAC determination measurements, each well containing 100 µL of sample, 10 acquisitions lasting 10 seconds were conducted, and the average values were plotted. DYNAMICS software implementing the Dynals algorithm was used for the data analysis**.**

Computational model: The RDKIT cheminformatics python library was used to extract 200 chemical descriptors from each molecule by feeding it with its SMILES string. Since the logCAC values were obtained from a mixture of monomer and mPEGMA (theoretical molar ratio: 90/10), descriptors for both the monomer and mPEGMA were obtained for each surfactant and their contributions were weighted by their molar ratio. The initial feature set were considerably reshaped by multiplying descriptors each other and by themselves: this was done to broaden the feature space and thus to have access to new combined descriptors. The resultant 20300 descriptors were then pruned to keep only those descriptors not excessively cross-correlated and with a good diversity. As far as concerns the first criterion, Pearson's r^2 of all features was computed to measure the extent of cross-correlation, while the criterion of diversity allows to identify those descriptors whose values do not change significantly across the samples. Ideally, no correlation should be observed between any descriptors, and all of them should assume as many different values as possible with a uniform frequency. Diversity can be expressed in many different ways, but the most common formula is Shannon's entropy. Thus, we decided to compute descriptor diversity as the ratio between the Shannon's entropy of the descriptor vector and the Shannon's entropy of the ideal descriptor vector if all its actual values were equally distributed. To compute diversity for a fictitious descriptor:

$$d_{real} = [1,1,0,0,0,0,0,0,0,0]$$

$$d_{ideal} = [1,1,1,1,1,0,0,0,0,0]$$

Where d_{real} is the descriptor vector with all its actual values, and d_{ideal} is the descriptor vector wherein all its values have the same frequency.

$$H_{d_{real}} = -\sum_{i=1}^{n} P(x_{real_i}) lnP(x_{real_i}) = -(0.2\ln(0.2)) - (0.8\ln(0.8)) = 0.321 + 0.178 = 0.499$$

$$H_{d_{ideal}} = -\sum_{i=1}^{n} P(x_{ideal_i}) lnP(x_{ideal_i}) = -(0.5\ln(0.5)) - (0.5\ln(0.5)) = 2 * 0.347 = 0.694$$

$$Diversity = \frac{H_{d_{real}}}{H_{d_{ideal}}} = \frac{0.499}{0.694} = 0.719$$

Where $P(x)$ is the frequency of the x-th value in the descriptor vector, $H_{d_{real}}$ is the Shannon's entropy of the descriptor vector, $H_{d_{real}}$ is the Shannon's entropy of the ideal descriptor vector, and D is the diversity associated to the descriptor vector. Diversity can assume any value between 0 and 1 and high values are associated to descriptor vectors with high amount of information.

The multicollinearity and diversity criteria were applied to narrow the feature space down to a more manageable number of descriptors. After setting both maximum tolerated cross-correlation and minimum accepted diversity to 0.9, number of features were reduced to 329. Then, we computed all the $\binom{329}{k}$ combinations, with k=1 and then k=2, to obtain a total number of 53956 different feature subsets, made of either one or two descriptors, which could either be a simple or combined (as described before). Since we had only 18 samples, value of k was deliberately kept small so that any resultant model could still be simple and interpretable. Once dataset was standardised through mean-centering, a linear regression model was trained on the dataset using one feature subset at a time, and model performance was assessed through the evaluation of

Pearson's r^2 and Root Mean Squared Error (RMSE). Then, a Leave-One-Out Cross-Validation (LOOCV) was carried out and its CV RMSE was also computed to validate model robustness. Finally, all models were sorted by the computed geometric mean between training and CV r^2 and first model was chosen. Cross-validation procedure is a common practice to validate results if dataset is such small that it cannot be split into training and test set. Thus, a common way to work around this is to split the dataset, which is made of n samples, into k folds so that a model can be trained on n-(n/k) samples and tested on n/k samples that have been left out of the model. Iteratively repeating the process and computing the mean of the desired performance metrics across all the iterations allows to have a better understanding of whether the model is robust or volatile. In our case, LOOCV is a particular case of k-fold cross-validation with k = n. Thus, for 18 times model was trained on 17 samples and RMSE of prediction was computed for the sample that was kept out.

S2. Supporting Figures and Tables

Table S2.1 Conversion, Mn, Đ and final copolymer ratio of the scale up reactions with CCTP with a target feed ratio of 90:10 mol/mol.

Entry	HbB:mPEGMA$_{300}$	% Conversion[a]	M_n[b] [KDa]	Đ[b]	Final HbM:PEG ratio[a] [mol/mol]
1	THFuA	60%	11.2	2.44	94:6
2	EA	63%	9.3	2.34	91:9
3	HPhOPA	77%	20	1.86	95:5
4	NMEMA	42%	7	1.46	94:6

[a)] % Conversion and final hydrophobic monomer:PEG ratios were calculated by ^{1}H-NMR; [b)] M_n and Đ were calculated by GPC.

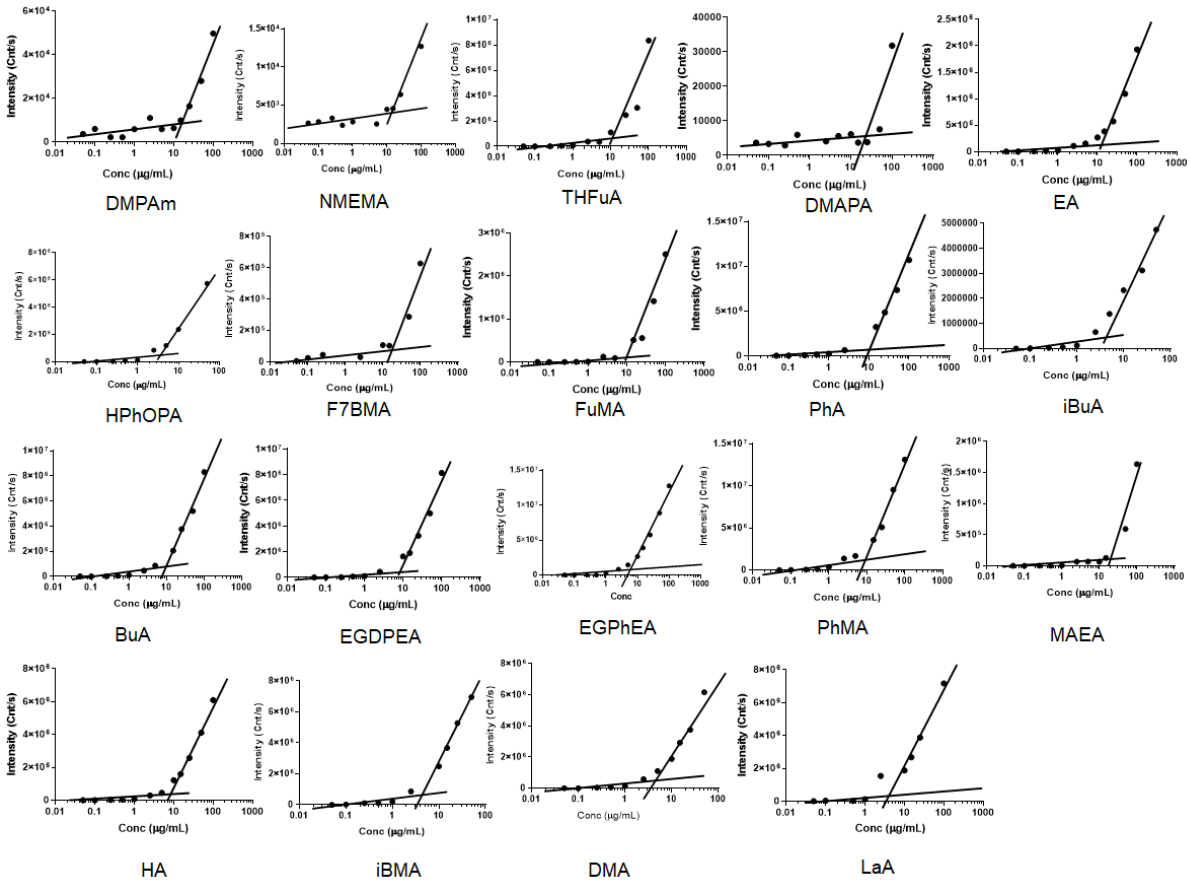


Figure S2.1 a) CAC of iBMA-co-mPEGMA$_{300}$, an example plot of the intensity of scattered light (kilo counts per second) as a function of concentration (µg/mL). The data showed that the scattering detected for the surfactant concentrations below the CMC is similar to deionized water. After the CMC was reached, the scattering intensity shows a linear increase with concentration. The intersection between the 2 lines, at 4.93 µg/mL, corresponds to the CAC/CMC. b) CAC plot of the intensity of scattered light (kilo counts per second) as a function of concentration (µg/mL) for all the 19 surfactants.

$$\frac{\sigma_{logCAC}}{RMSE_{model}} = \frac{0.233}{0.074} = 3.027$$

Equation S2.1 Relationship between the standard deviation of the logCAC (σ_{logCAC}) and the $RMSE_{model}$