Chemical Science

Accepted Manuscript

This article can be cited before page numbers have been issued, to do this please use: K. Sharma, A. McCorry, S. Boobier, J. Mottram, R. Napier, I. W. Ashworth, J. Blacker, N. Kapur, S. Warriner, M. H. Wright and B. N. Nguyen, *Chem. Sci.*, 2024, DOI: 10.1039/D3SC06311A.

ROYAL SOCIETY OF CHEMISTRY

rsc.li/chemical-science

View Article Online
DOI: 10.1039/D3SC06311A

ARTICLE

Activation of Fluoride Anion as Nucleophile in Water with Data-Guided Surfactant Selection

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

DOI: 10.1039/x0xx00000x

Krishna Sharma,[a] Alison McCorry,[a] Samuel Boobier,[a] James Mottram,[a] Rachel Napier,[a] Ian W. Ashworth,[b] A. John Blacker,[a] Nikil Kapur,[c] Stuart L. Warriner,[a] Megan H. Wright,[a] and Bao N. Nguyen *[a]

A principal component surfactant_map was developed for 91 commonly accessible surfactants for use in surfactant-enabled organic reactions in water, an important approach for sustainable chemical processes. This map was built using 22 experimental and theoretical descriptors relevant to the physicochemical nature of these surfactant-enabled reactions, and advanced principal component analysis algorithms. It is comprised of all classes of surfactants, i.e. cationic, anionic, zwitterionic and neutral surfactants, including designer surfactants. The value of this surfactant_map was demonstrated in activating simple inorganic fluoride salts as effective nucleophiles in water, with the right surfactant. This led to the rapid development (screening 13-15 surfactants) of two fluorination reactions for β-bromosulfides and sulfonyl chlorides in water. The latter was demonstrated in generating a sulfonyl fluoride with sufficient purity for direct use in label-ling of chymotrypsin, under physiological conditions.

Introduction

Surfactant-enabled organic reactions in water were first reported in the 1970s,[1,2] wherein the surfactants were thought to stabilise organic emulsions and the charged hydrophilic terminus may have a role in stabilizing the reaction transition states.[3] ☐Due to the sustainable nature of water as a reaction medium, there has been a resurgence of interest in these reactions,[3–6] particularly in the use of neutral designer surfactants.☐[4] These surfactants have been effectively implemented in supporting various organic and cross-coupling reactions,[7–9] including those which employ reactive reagents which may rapidly decompose in water.[10] ☐Some of these reactions are shown in Figure 1a.[11,12] However, fundamental understanding of how these surfactants work and how to rationally select suitable surfactants from the hundreds of commercially available surfactants remains a key challenge in the field, due to the complex nature of the reaction mixtures.[13–17]

In this paper, we report the first surfactant map (***surfactant_map***) constructed for organic reactions. It was built by Principal Component Analysis (PCA) of surfactant properties, which were selected based on physicochemical understanding of the surfactant-enabled reactions in water. It allows rapid and rational selection of the optimal surfactant for any organic reaction, which in many cases are not designer surfactants. We demonstrated this powerful data-based approach by employing the ***surfactant_map*** to turn on the nucleophilicity of fluoride in water. Fluoride anion is well-known to be extensively hydrated in water, suppressing its ability to act as a nucleophile in competition with water or hydroxide anion.[18] On the other hand, fluoride salts are often poorly soluble in organic solvents, necessitating the use of elaborate fluorination agents, such as $Me_4NF.t$AmylOH,[19] at high temperature or the highly corrosive KHF_2 for nucleophilic fluorination.[20,21] The use of a surfactant allows fine tuning of the fluoride anion hydration at the water-organic interface, successfully revealing the nucleophilicity of fluorides in water and enabling nucleophilic fluorination reactions using simple, relatively safe and readily available $KF.2H_2O$ as the source of fluoride anion.

[a] School of Chemistry, University of Leeds, Woodhouse Lane, LS2 9JT, United Kingdom.
[b] Chemical Development, Pharmaceutical, Technology and Development Operations, AstraZeneca, Macclesfield, SK10 2NA, United Kingdom.
[c] School of Mechanical Engineering, University of Leeds, Woodhouse Lane, LS2 9JT, United Kingdom.
Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 10.1039/x0xx00000x

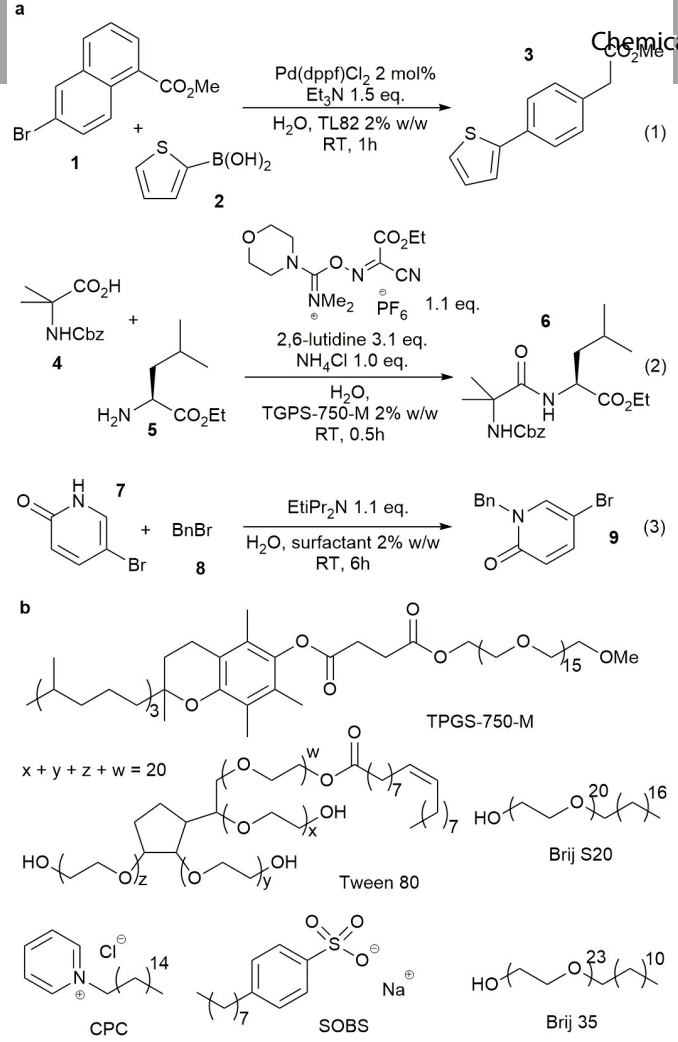


Figure 1. Example of organic reactions in water enabled by surfactants, (a) Representative reactions enabled by traditional and designer surfactants, TL82 is a mixture of Tween 80/lecithin 8:2, (b) Representatives of different classes of surfactants.

Results and discussion

Surfactant map for organic reactions

The successes of designer surfactants in enabling organic reactions in aqueous media, with or without a small amount of organic solvent to overcome solubility limits,[3,4] can potentially lead to many sustainable chemical manufacturing processes. This minimises the use of volatile and toxic organic solvents, particularly the difficult to replace dipolar aprotic solvents. In addition, the surfactants may modify the reactivity of organic compounds and reagents at the organic-water interface,[25] giving rise to new reactivity which may not be accessible otherwise. The relative high cost and limited supply of designer surfactants, compared to the many readily available commercial, food-grade surfactants, is currently a barrier for their wider adoption. On the other hand, many researchers have established that other surfactants can work just as well, and sometimes better than designer surfactants. [13,26,27] In this study we aim to develop a data-based tool for rational, rapid surfactant selection, from a pool of 100 common and commercially available surfactants for organic reactions. This tool, which we call ***surfactant_map***, is based on the proven PCA approach utilised in solvent selection for organic reactions and ligand selection in catalysis.[28–31] Due to the length of the study, the latest designer surfactants, i.e. Savie,[32] and APGS 2000-M,[33] are not included but will be in the next iteration of the map.

In order to capture the relevant data on surfactant-enabled reactions, the physicochemical aspects of these reactions were investigated. Whilst this type of reaction is colloquially referred to as 'micellar catalysis', the nature of the system can be complex, particularly at higher effective organic concentration (>0.2 M) and in the presence of commonly employed inorganic bases, which increase the ionic strength of the aqueous phase and affect concentrations of species at emulsion surface (Figure 2a). Dynamic light scattering measurements of surfactant TPGS-750-M in water (0.5% w/w) showed that the system moved away from being micelles at 0.2 M loading of toluene, accompanied by visible change to a milky emulsion. These were confirmed by optical microscopic measurements, which gave a distribution of organic droplet size in the range of 0.2-6 microns O.D. A similar droplet size distribution was observed in the reaction (3) (Figure 2b), performed with surfactant TPGS-750-M, with increasing aggregation of droplets as the reaction proceeded.

Another uncertainty is the partition of the reactants between the organic (inside the emulsions) and aqueous phases.[34] While this has an obvious impact on reaction rate, the partition is highly dependent on organic compounds and surfactant. The extent of partition, and where the reaction happens, are critical pieces of information to explain how the surfactant influences the reaction transition state and enables the transformation. Thus, ^{1}H NMR experiments were performed to evaluate the partition of the starting materials and product of reaction (3) in the presence of a surfactant in D_2O by comparing their chemical shifts with the corresponding peaks in D_2O or d_{12}-cyclohexane (Figure 2c, Figure S29-S46). Reaction (3) was selected due to the two tautomeric forms of starting material **7** with very different chemical shifts in water and organic solvents. Six surfactants, TPGS-750-M, Tween 80, Brij S-20, Brij 35, CPC and SOBS (Figure 1b), of different classes, e.g. neutral, cationic and anionic, were studied at 2.7% w/w loading in D_2O. Starting material **7** was consistently found to be in an aqueous-like environment (except with SOBS), while starting material **8** was in an organic-like environment. The partition of product **9** was less clear-cut and was surfactant dependent (Section 1.3.3 of supporting information). In TPGS-750-M solutions, the chemical shifts of **9** suggested a dynamic exchange between two environments, i.e. organic and aqueous. The chemical shifts of both **7** and **9** change in small degrees when the amount of surfactant was doubled, indicating observable interaction between each compound and the surfactant molecules (Section 1.3 of supporting information). In addition, DOSY experiments showed different changes in diffusion coefficients of **7** (6.3×10^{-10} m^2.s^{-1} to 4.5×10^{-10} m^2.s^{-1}) and 8 (4.0×10^{-10} m^2.s^{-1} to 4.1×10^{-11} m^2.s^{-1}) in D_2O upon inclusion of 2.7% w/w of TPGS-750-M. These suggested inclusion of **8** inside the TPGS-750-M micelles/emulsions, and weak interaction of **7** with the interface of these micelles/emulsions, in agreement with the changes in ^{1}H NMR chemical shifts above. Similar DOSY interaction between a catalyst and anionic surfactants was also

View Article Online
DOI: 10.1039/D3SC06311A

observed by Scarso and Strukul in a surfactant-enabled Baeyer-Villiger reaction.[35] Taken all together, the NMR evidence suggested that the reaction likely happened at the organic-water interface, assisted by the surfactant.

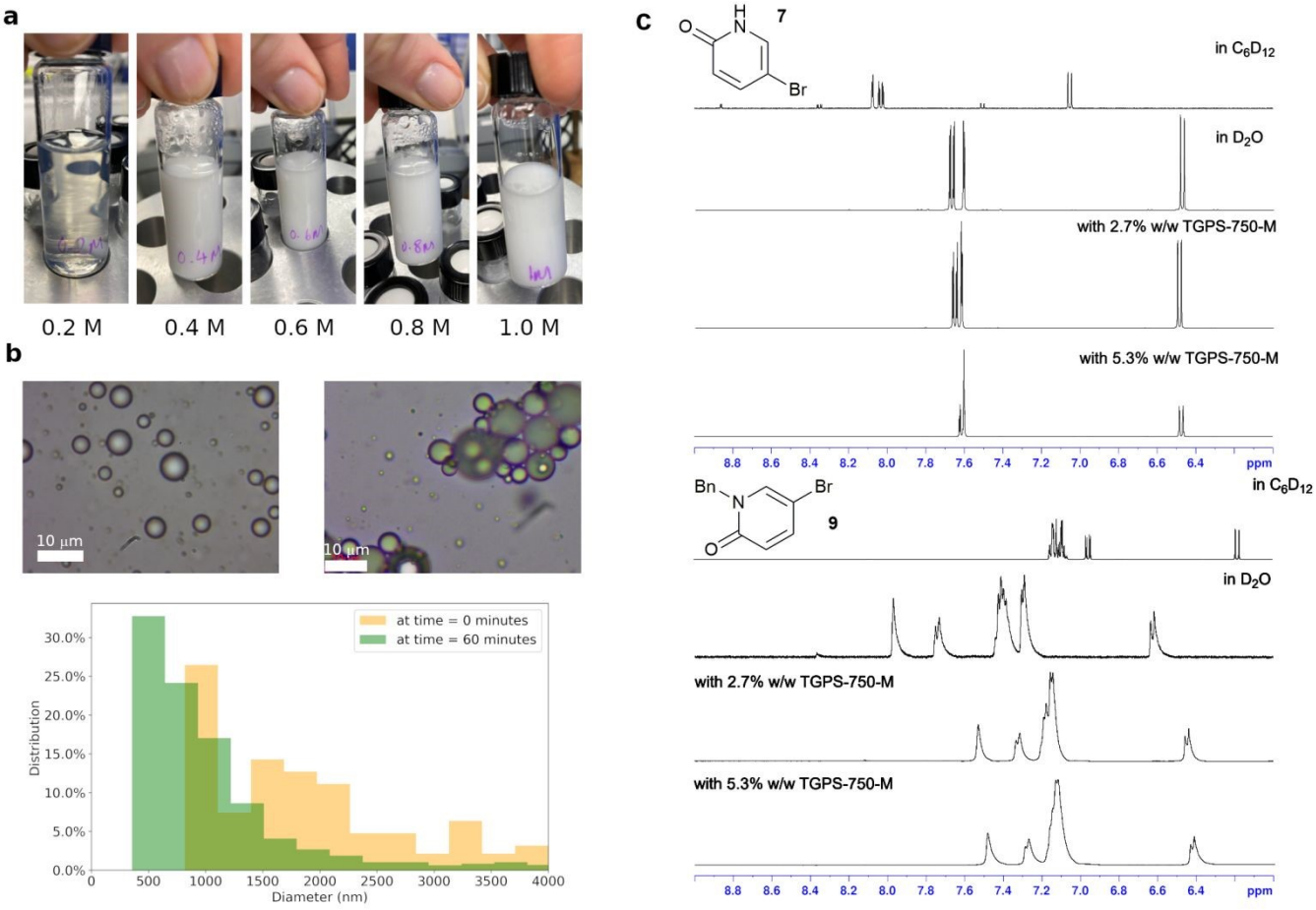


Figure 2. Physicochemical behaviours of surfactant-enabled reactions and the data-based ***surfactant_map***, (a) Visual observations of TPGS-750-M 2% w/w in water at 0.2, 0.4, 0.6, 0.8 and 1.0M effective toluene loading, (b) Microscope images of toluene in TPGS-750-M 2% w/w and reaction mixture (3) at 0 and 60 minutes, and their particle size distributions, (c) Chemical shifts of **7** and **9** in D_2O, d_{12}-cyclohexane and D_2O in the presence of a surfactant.

Based on the observations above, a number of descriptors, which represent the computational properties of the surfactant molecule (2D and 3D structural information, e.g. SASA and number of double bonds, divided into hydrophobic and hydrophilic fragments (Figure 3) to reflect emulsion properties by hydrophilic-lipophilic balance),[36,37] its interaction with water (number of OH groups, e.g. Brij S20 (Figure 1b), ΔGsolv, HOMO and LUMO energies, etc.), its micellar properties (critical micelle concentration, aggregation number and micelle size), and its emulsion properties (zeta potential, contact angle, and hydrophilic lipophilic balance), were selected to build the ***surfactant_map*** for 100 common surfactants (Table S13, selected based solely on commercial availability in order to expand the pool of surfactants for synthetic reactions). While the reaction mixtures are often emulsions, these micellar properties do represent surfactant-surfactant interactions, which are relevant to emulsions. The micellar properties were curated from the literature, while the emulsion properties were measured experimentally. The computational descriptors were derived using either the *rdkit* cheminformatics package or PM6 molecular modelling calculation. One of the most difficult aspects of numerically representing surfactants in emulsions/micelles is how to represent the charge of the surfactant species, which is a discrete rather than continuous variable, without artificially clustering cationic, anionic and neutral surfactants in the ***surfactant_map***. Thus, the intrinsic Hirshfeld charge of the atom with the most negative charge of the hydrophilic end of the surfactant molecule was used for each surfactant instead.[38] While the majority of the hydrophobic chains are linear alkyl chains, a small number of them include cis double bonds, which can have significant impact on the packing of the surfactant molecules in micelles and emulsions, and consequently the dynamic rate of material exchange at the organic/water interface. Therefore, a descriptor for the number of cis double bonds in the hydrophobic part of each surfactant was included in the dataset. The full list of descriptors is summarised in Table 1.

View Article Online
DOI: 10.1039/D3SC06711A

Table 1. Descriptors for *surfactant_map* and their sources

Descriptor	Representing
Critical micelle concentration[a]	Surfactant-surfactant interactions
Aggregation number range[a]	Surfactant-surfactant interactions
Micelle size range[a]	Surfactant-surfactant interactions, homogeneity
Contact angles (left and right)[b]	Surface tension and wettability
Zeta potential[b]	Charge environment around micelles, emulsions
Hydrophilic Lipophilic Balance (HLB)	Emulsion stability
Hydrophilic fragment rotatable bonds[c]	Flexibility of surfactant molecules and emulsion flexibility/stability
Hydrophobic fragment rotatable bonds[c]	Flexibility of surfactant molecules and emulsion flexibility/stability
Hydrophilic fragment longest chain length[c]	Size of the interface layer between organic and aqueous phases
Hydrophobic fragment longest chain length[c]	Capability for stabilizing organic phase inside emulsions
Hydrophilic fragment volume[c]	Packing of surfactant and stability of emulsion
Hydrophobic fragment volume[c]	Packing of surfactant and stability of emulsion
Hydrophilic fragment surface area[c]	Packing of surfactant and stability of emulsion
Hydrophobic fragment surface area[c]	Packing of surfactant and stability of emulsion
Hydrophobic fragment number of C=C bonds	Flexibility of surfactant molecules and emulsion flexibility/stability
Hydrophilic fragment number of OH groups	Capability of H-bonding at the emulsion interface
Hydrophilic fragment ΔG_{solv}[c]	Stability of emulsion
Hydrophilic fragment dipole moment[c]	Stability of emulsion
Hydrophilic fragment HOMO energy[c]	H-bonding capability and interactions with transition states
Hydrophilic fragment LUMO energy[c]	H-bonding capability and interactions with transition states
Hydrophobic fragment dipole moment[c]	Stability of emulsion
Hirshfeld charge for most negative heteroatom[c]	Interactions with transition states

Micellar properties are in **blue**; emulsion properties are in **red**; and molecular properties are in **black**; Sources of data: [a]Experimental (literature); [b]Experimental (measured); [c]*rdkit,* PM6.

Whilst experimental data are highly valuable, the inclusion of experimental descriptors led to a significant amount of missing data, i.e. 263 entries, accounting for 10.5% of the total 2500 (100 surfactants × 25 descriptors). This prevents the use of a standard PCA algorithm. Removing experimental data completely risks losing reliable information on intermolecular interaction between surfactant molecules and with solvent. Consequently, four modified-PCA algorithms: PPCA (estimate/impute missing values with Gaussian probability),[39] BPCA (estimate/impute missing values with Bayesian probability and expectation-maximization repetitive algorithm),[40] NLPCA (artificial neural networks which convert a dataset to principal components and reconstruct it to impute missing data),[41] and NIPALS (iterative method which skips the missing data, see Section 2.5.1 of supporting information for more complete explanation of the algorithms),[42] which have been demonstrated to work with varying degrees of missing data, were compared based on their capture of data variance with up to 5 principal components (PCs, Figure 4a). BPCA did not cope well with the missing data. PPCA gave the best result with R^2 = 0.89 for 5 PCs, but very poor capture of variation in the first 4 PCs. For visualization purposes, NLPCA was the best method, achieving R^2 = 0.78 with only 3 PCs, and NIPALS was the next best choice. However, the PCs generated by NLPCA were found to vary significantly between different runs (the difference in PC1 (0.01±0.13 over 100 surfactants) was -0.176 to 2.317 between 2 runs). This was due to due to the reconstruction stage of the algorithm and the large amount of missing experimental data for 9 surfactants, which consistently lack >5 experimental descriptors, leading to randomization of the produced PCs. Varying the number of cycles of optimization/imputation between 100 and 10000 did not improve the reproducibility, and averaging 100 different runs did not produce sensible results, i.e. most surfactants were pushed close together in 3D space. This is a known issue with NLPCA algorithm, which works best with randomly missing data instead of large amount of missing data on certain rows.[43] On the other hand, NIPALS algorithm produced identical PCs for each surfactants in repeated runs on all 100 surfactants. Careful examination of the 9 surfactants with significant missing experimental data showed that they were placed close to surfactants with which they share little similarity within the 3D map with PC1-3. Consequently, these 9 surfactants were removed, leaving 91 surfactants for the final analysis, and reducing the percentage of missing data from 10.6% to 8.8% (Table S14). A 3D map based on 3 PCs generated with NIPALS for these 91 surfactants was built and shown in Figure 4b as our ***surfactant_map***.

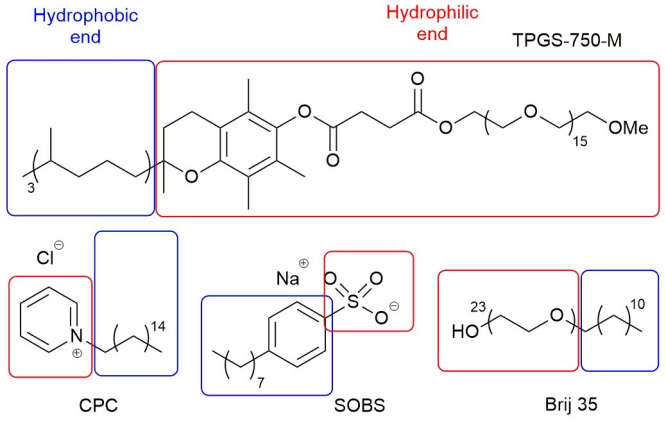


Figure 3. Example of hydrophobic and hydrophilic fragments for surfactants.

This ***surfactant_map*** showed some degree of the expected clustering of cationic, anionic and zwitterionic surfactants. Crucially, the most numerous neutral surfactants (blue) occupy the most space in the PCA map, mixing well with the anionic

(orange) and zwitterionic surfactants (green) but not the cationic surfactants (red). Zwitterionic surfactants and cationic surfactants are positioned close to others in the same class, suggesting that they behave similarly to each other. Analysis of the map suggests that PC3 is significantly influenced by nominal charge of the surfactant molecule; cationic surfactants have higher PC3 values and zwitterionic surfactants have lower PC3 values; and neutral surfactants cover a very wide range of PC1 and PC2 values. These are consistent with the loadings for PC1-3 (Table S15), wherein zeta potential, solvation energy and the most negative atomic charge on the hydrophilic fragment contribute the most to PC3. The most significant contributors to PC1 are centred on the volume, surface area and flexibility of the hydrophobic and hydrophilic fragments of each surfactant, i.e. hydrophobic/ hydrophilic balance. PC2 was mainly derived from a combination of contact angles, zeta potential and volume/area/flexibility properties of the hydrophobic fragment. Importantly, the designer surfactants (purple) occupy a relatively small, but central portion of surfactant space compared to the rest of the surfactants (Figure 4c). This explains the relative success of these surfactants in surfactant-enabled reactions and, at the same time, highlights the risk of exclusively focusing on designer surfactants when screening surfactants for a given reaction. A small number of outliers, e.g. Croduret 25-LQ, and Croduret 50-SS, can be attributed to their unique structures and properties, i.e. PEG-25 or PEG-40 hydrogenated castor oil. Interestingly, the surfactant IGEPAL CO-720, which is structurally very similar to IGEPAL CA-720 and Triton-X-102, was placed closest to these surfactants in the map, showcasing the ability of NIPALS algorithm to cope with missing data for IGEPAL CA-720 and Triton-X-102, if there is significant similarity in the remaining descriptors between surfactants. The ***surfactant_map*** and workflow led to rapid identification of sulfobetain-16 and TTAB as the best surfactants for reaction (3), giving 95% yield (5% of *O*-alkylated product) in 20 minutes at 45 °C, in comparison to 64% yield with TPGS-750-M (Figure 5). A similar surfactant screen for the reaction between **7** and allyl bromide, instead of benzyl bromide, also identified sulfobetain-16 and TTAB as the optimal surfactant (100% for *N*-allylated product with sulfobetain-16). Nevertheless, the true usefulness of this ***surfactant_map*** needs to be demonstrated in enabling new chemical reactions.

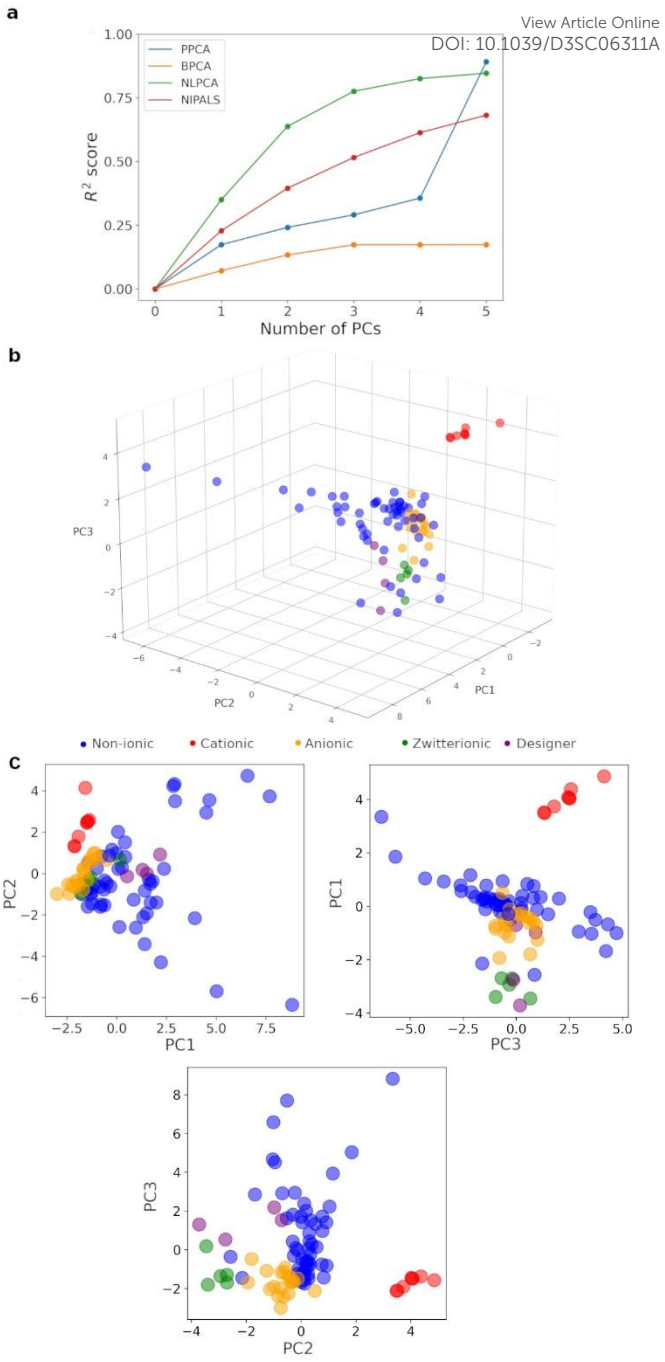


Figure 4. (a) Variance captured with different PCA algorithms *vs* number of PCs. (b) NIPALS-derived 3D ***surfactant_map***. (c) 2D projections of the ***surfactant_map*** showing the information in each PC.

Nucleophilic fluorination with fluoride anion

While the use of fluoride as a nucleophile in mixture of organic solvents and water have been reported in recent years,[44,45] often at elevated temperature, reactions in water alone are rare due to the very strong hydration of fluoride anion in water. The only example without a co-solvent in the literature employs KHF_2, which is corrosive against glass vials and reactors, as the source of fluoride.[20] While the nucleophilicity of fluoride anion in organic solvent is much improved, the high lattice energy of readily available inorganic fluorides means that their solubility

is often very low in these solvents. Previous solutions for this conundrum included the use of phase transfer catalyst and water/organic solvent combinations,[46–48] and elaborate fluoride reagent such as NMe_4F and $NMe_4F.t$Amyl-OH,[19] which have better solubility in organic solvents, or SuFEx reagents,[49] which generate *in situ* fluoride anion in organic solvents. An alternative, yet unexplored, solution is to modify the environment around the fluoride anion, and thus its reactivity, at the organic/water interface through the use of surfactant (Figure 6a). This has the benefit of circumventing the solubility limitation, while fine tuning the nucleophilicity and basicity of fluoride anion to achieve selective reactions. Thus, the ***surfactant_map*** was used to select the surfactants which can activate fluoride anion as a nucleophile in water for synthesis. To achieve this, a selection of 8-10 surfactants (***screen1***, sufficient to cover most areas of the map) were made using a Python script which randomises the selection of surfactants while maximizing their coverage of the ***surfactant_map***. The origin point can be set for the first hit surfactant which enables reaction, if known, or TPGS-750-M, which locates approximately in the center of the surfactant space. The best results obtained with screen1 will lead to the selection of an additional 5-6 surfactants in the region around the best surfactants for ***screen2***, leading to the optimised surfactant for a given reaction. This workflow is described in Figure 5.

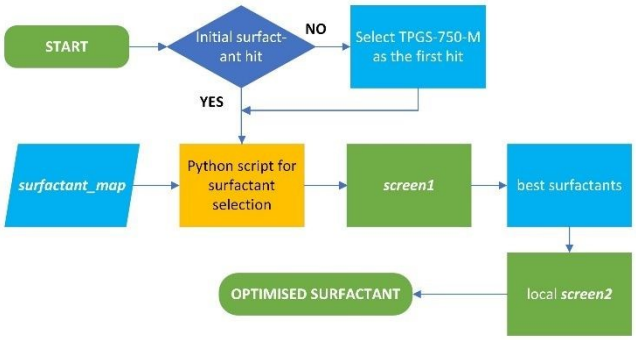


Figure 5. Workflow employing the ***surfactant_map*** for optimising surfactant-enabled reactions.

This workflow was first demonstrated with the fluorination of episulfoniums, generated *in situ* from β-bromosulfides (4), reported by Gouverneur and co-workers,[47] and the results are summarised in Figure 6. A non-chiral catalyst, 1,3-bis[3,5-bis(trifluoromethyl)phen-yl]urea **14**, was used for simplicity, and the competition between the fluoride anion and water as nucleophiles, giving products **11** and **12**, was used to evaluate the effectiveness of the surfactant in improving the reactivity of the fluoride anion (Figure 6b). Initial reaction using 3 equivalents of CsF in a water/toluene (9:1) mixture without a surfactant at room temperature gave little conversion and an unexpected product, alkene **13** (entry 1, Table 2). Formation of alkene via nucleophilic attack of a fluoride anion on the S atom of an episulfonium, instead of the C atoms, has previously been reported by Helmkamp.[50] Changing the organic solvent to the more polar 1,2-dichloroethane (DCE) led to improved reactivity, although not necessarily a significant change in products ratio (entry 2). The best results, 57%, of **11** and the best combined yield for products of fluoride nucleophilic attack on the C and S atoms, i.e. 92% for **11** and **13**, was obtained with Span 80 as the surfactant in ***screen1***, starting from Span 20 as the first hit surfactant and performed with gentle heating (50 °C). This was followed by ***screen2*** using Span 85, Span 60, Span 40 and Tween 85, which are near Span 80 in the ***surfactant_map***. Nevertheless, Span 80 remained the best performing surfactant. The surfactants had a clear impact on the distribution of products between hydrolysis (**12**), nucleophilic attack on C (**11**) and on S (**13**) (Table S18 of the supporting information). More hydrolysis product **12** was observed with surfactants with shorter hydrophobic chain lengths (C10-12: SDS, lauryl betaine, PS-750-M and Brij 35). Selectivity between **11** and **13** is relatively poor, with high yields for **13** often associated with high yields of the desired product **11**. The lowest **11**:**13** ratios were observed lauryl betaine and Brij 52, straight chain surfactants with short hydrophilic fragment. Further optimization enabled replacement of CsF with the much cheaper $KF.2H_2O$ (6 eq.) without significant change to reaction performance. Attempts at replacing DCE with toluene led to a lower selectivity, likely via changes to the stability of the emulsions, giving 40% of the desired product **11** (entry 9). The standard reaction conditions, using Span 80 and 6.0 eq. of $KF.2H_2O$ were successfully applied to three derivatives of **10** as substrates, giving moderate to good yields of the desired fluorides (Figure 6d).

To rule out a simple solubility effect, the reaction was performed without surfactant, using either tetrabutylammonium fluoride (TBAF) at room temperature or a combination of $KF.2H_2O$/dibenzo-18-crown-6 ether (entry 10 and 11). Both conditions increase the concentration of fluoride anion in the organic phase. No product **11** was observed with TBAF, while a 1:3.13 ratio of **11**:**12** was observed with $KF.2H_2O$/dibenzo-18-crown-6 ether. These data support our hypothesis of reactivity attenuation at the water/organic interface by surfactant.

Sulfonyl fluorides have been widely employed as reactive probes in chemical biology, thanks to their biocompatibility and specific reactivity toward reactive serine, threonine, lysine, tyrosine cysteine and histidine residues.[51,52] Their synthesis often involves activation of sulfonamides, deoxygenation of sulfonic acids or electrochemical oxidation of thiols under harsh or highly reactive conditions.[51] Direct fluorination of sulfonyl chlorides under aqueous conditions is highly desirable, as the sulfonyl fluorides may be readily used in peptide labelling without the need for purification with flash chromatography. This has been shown possible with the help of a phase transfer catalyst or in a combination of MeCN and water, albeit using the highly corrosive KHF_2 reagent.[20,21,49] Given our success with fluorination of β-bromosulfides, we hypothesised that the use of a surfactant will enable the use of a simple fluoride salt. Thus, a surfactant screen (***screen1***, using TPGS-750-M as the starting point) using 10 surfactants was carried out for reaction (5) from *p*-toluenesulfonyl chloride with 1.5 eq. of $KF.2H_2O$ at room temperature in water (Figure 7a). The best result was obtained with CTAC, giving 96% yield after 3 hours, compared to 13% yield without any surfactant (Table 3, entries 1 and 11). Thus,

the surfactant space around CTAC was explored in ***screen2*** with CPC, CTAB, DDAB and DTAB (entries 3-7).

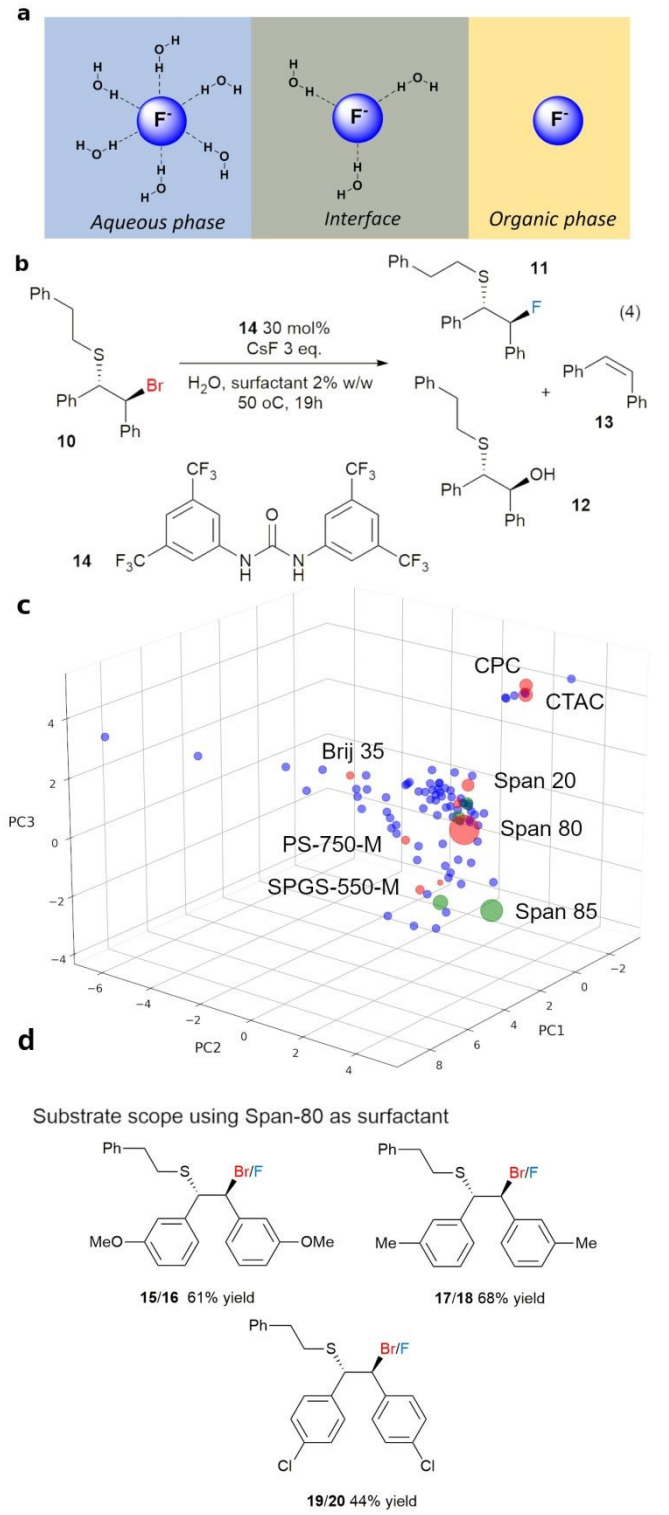
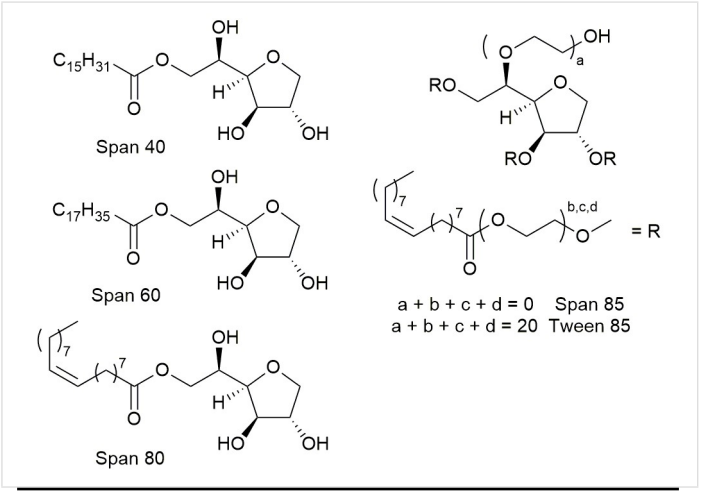


Figure 6. (a) Modification of the nucleophilicity of fluoride anion at the organic/water interface, (b) fluorination of β-bromosulfides in water with CsF and surfactant in ***screen1***, (c) results from ***screen1*** (red) and ***screen2*** (green) for reaction (4) showing the local area with best selectivity, after 19 hours at 50 °C, the tested surfactants are in red, and the size of the marker represent the **11**:**12** ratio, (d) substrate scope and ^{1}H NMR yields.

Table 2. Surfactant optimization for fluorination of β-bromosulfide **10**

View Article Online
DOI: 10.1039/D3SC06311A



No.	Solvent	Fluoride	Surfactant	**10**:**11**:**12**:**13**[a]	**11**+**13** (%)
1[b,c]	Toluene:H_2O (10:90)	CsF	None	81:**6**:7:6	12
2[b,c]	DCE:H_2O (10:90)	CsF	None	69:**10**:9:12	22
3	DCE:H_2O (10:90)	CsF	Span 80	0:**57**:8:35	92
4	DCE:H_2O (10:90)	CsF	Span 85	0:**52**:13:35	87
5	DCE:H_2O (10:90)	CsF	Span 40	0:**38**:32:30	68
6	DCE:H_2O (10:90)	CsF	Span 60	0:**39**:31:30	69
7	DCE:H_2O (10:90)	CsF	Tween 85	0:**44**:25:31	75
8[d]	DCE:H_2O (10:90)	KF.2H_2O	Span 80	0:**58**:13:29	87
9	Toluene:H_2O (10:90)	KF.2H_2O	Span 80	0:**40**:28:32	72
10[c]	DCE:H_2O (10:90)	TBAF	None	75:**0**:6:18	18
11[d]	DCE:H_2O (10:90)	KF.2H_2O	Dibenzo-18-crown-6	18:**16**:50:16	32

Standard conditions [**10**] = 0.15M, 3.0 eq. of fluoride, 30 mol% of catalyst 14, 2% w/w surfactant in water, 19h at 50 °C; [a] Determined by ^{1}H NMR; [b] [**10**] = 0.08M; [c] at room temperature; [d] 6.0 eq. of KF.2H_2O. The full list of surfactants and experiments is in Table S17.

DOI: 10.1039/D3SC06311A

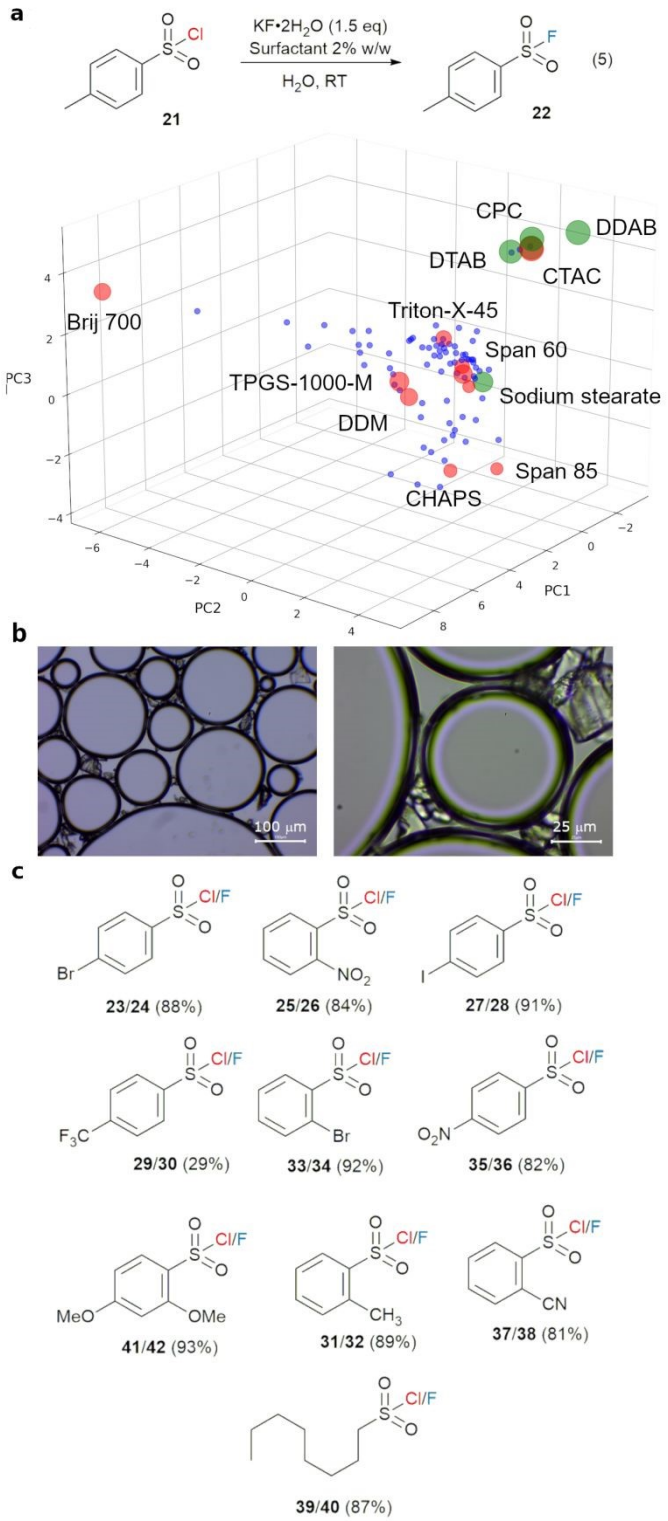
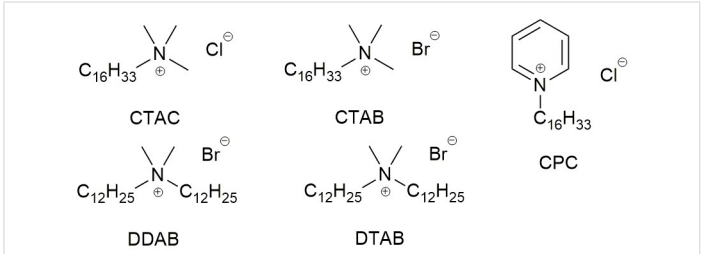


Figure 7. (a) surfactant screen results for reaction (5) showing the local area with best reaction yield after 3 hours at room temperature, the tested surfactants from ***screen1*** are in red and from ***screen2*** are in green, and the size of the marker represents the yield of **22**, (b) microscope images of reaction mixtures without KF.2H_2O; c, substrate scope and yield of fluorinations of sulfonyl chlorides in water with KF (3.0 eq.) and CTAC as surfactant.

Table 3. Surfactant *screen2* for fluorination of 21 to 22[a]



No.	Surfactant	^{1}H NMR yield of **22** (%)
1	None	13
2	Span 60	57
3	Span 80	25
4	Span 85	25
5	Triton-X-45	41
6	Brij-700	46
7	TPGS-1000-M	59
8	1-Dodecanesulfonic acid sodium salt	35
9	PS-750-M	49
10	CHAPS	29
11	CTAC	96
12	CPC	94
13	CTAB	90
14	DDAB	96
15	DTAB	88
16	Sodium stearate	65
17	18-crown-6	77

[a] 0.15 M substrate, 1.5 eq. of KF.2H_2O, 2% w/w surfactant at room temperature, 3h reaction time.

The best performing class of surfactant are cationic surfactants, which consistently gave high yields (88-96%) of **22**. CTAC and DDAB are the best performing surfactants, both giving 96% yield of **22**. As these cationic surfactants can also act as phase transfer catalysts, the reaction mixtures of **21** and CTAC (without KF.2H_2O) were examined with a microscope. Clear emulsions (50-500 μm), which resisted pressure, were observed. Surprisingly, compound **21** persisted as observable crystals in this system, instead of being completely dissolved as an organic phase inside the emulsions (Figure 7b). The lack of an organic phase, and the large size of the emulsions suggested that these are bilayer emulsions, representing a completely different reaction system compared to the standard view of 'micellar catalysis' (Section 5.3 of supporting information). Strong interfacial interaction between these crystals and the emulsions led to clustering, and a fluorination reaction which happens at the interface of the emulsions, crystalline **21** and the aqueous phase. To separate the roles of the surfactant in stabilising emulsions and facilitating phase transfer, 18-crown-6, a known phase transfer catalyst for potassium cation, was used instead. This led to a good conversion of 77%, highlighting the dual nature of the surfactant in this reaction.[53] As CTAC is significantly cheaper and more readily available than DDAB, it was selected as the surfactant of choice for substrate scope study (Figure 7c). Excellent isolated yields (81-93%), via a simple

View Article Online
DOI: 10.1039/D3SC06711A

filtration with most aryl sulfonyl chlorides (**23**-**37**), were obtained with a wide range of aryl and alkyl sulfonyl chlorides. ^{1}H and ^{19}F NMR spectra of the crude reaction mixture indicated complete conversion in most cases, with partial solubility of the sulfonyl fluorides in water being the reason behind imperfect isolation yields in some cases. The reaction tolerates a wide range of substituents on the phenyl group at *o*- and *p*-positions. The only poor yield was obtained with a strongly electron-withdrawing *p*-CF_3 substituent, with complete consumption of the sulfonyl chloride. This can be attributed to the low stability of the electron-poor sulfonyl fluorides against hydrolysis.[54]

The effectiveness of this new synthetic protocol for sulfonyl fluorides was demonstrated in a labelling experiment with the serine hydrolase chymotrypsin (Figure 8). Chymotrypsin is known to react with sulfonyl fluoride-containing protease inhibitor through the reactive active site serine.[55,56] 4-(Prop-2-yn-1-yloxy)-benzene-1-sulfonyl chloride (**43**) was chosen as the starting material for conversion to sulfonyl fluoride **44** at room temperature. Substrate **43** is an active sulfonyl chloride, which can readily react with methyl glycolate in the presence of a base at 0 °C, and is thus susceptible to hydrolysis.[57] It also contains an alkyne functional group which allow for further functionalization via 'click' reaction. Thus, it presents a good test for fluorination vs hydrolysis with a complex substrate. Reaction between **43** and $KF.2H_2O$ was carried out using the optimised conditions above, and the product was purified by extraction with DCM. Once isolated, one equivalent of **44** in DMSO (calculated based on 100% conversion of **43**) was added to a solution of bovine chymotrypsin in phosphate buffer pH 7.4 at a final concentration of 10 μM. LC-MS of the reaction mixture was compared to a control sample without **44** at 1 and 20 hours (Figure 8b). MS data showed a significant new set of peaks corresponding to a single labelling event, i.e. $[M+195]^+$ for $[chymotrypsin+C_9H_7O_3S]^+$, even after just 1 hour. At 20 hours, the labelling was judged complete.

Conclusions

In conclusion, we demonstrated that surfactant-enabled organic reactions in water are complex and diverse physiochemical systems, in which the desired reaction often occurs at the interface and can be influenced by the choice of surfactants. Given this complexity, the ***surfactant_map*** we developed, based on physicochemical understanding of these emulsified systems, can enable rapid screening of surfactants which enable organic reactions in water. The effectiveness of the map was demonstrated in fine-tuning the reactivity of simple fluoride salts at the organic-water interface, enabling fluoride anion as a nucleophile in reactions with β-bromosulfides and sulfonyl chlorides while suppressing hydrolysis of the substrate as side reactions. These outcomes were demonstrated in the simple synthesis of a complex sulfonyl fluoride and subsequent successful labelling of chymotrypsin. Work to expand surfactant-enabled nucleophilic fluorinations with simple fluorides is on-going in our group and will be disseminated in due course. Importantly, this ***surfactant_map*** underpins a rational and data-based approach to surfactant selection in surfactant-enabled organic reactions in water, a rapidly growing area of green chemistry in both academia and industry.

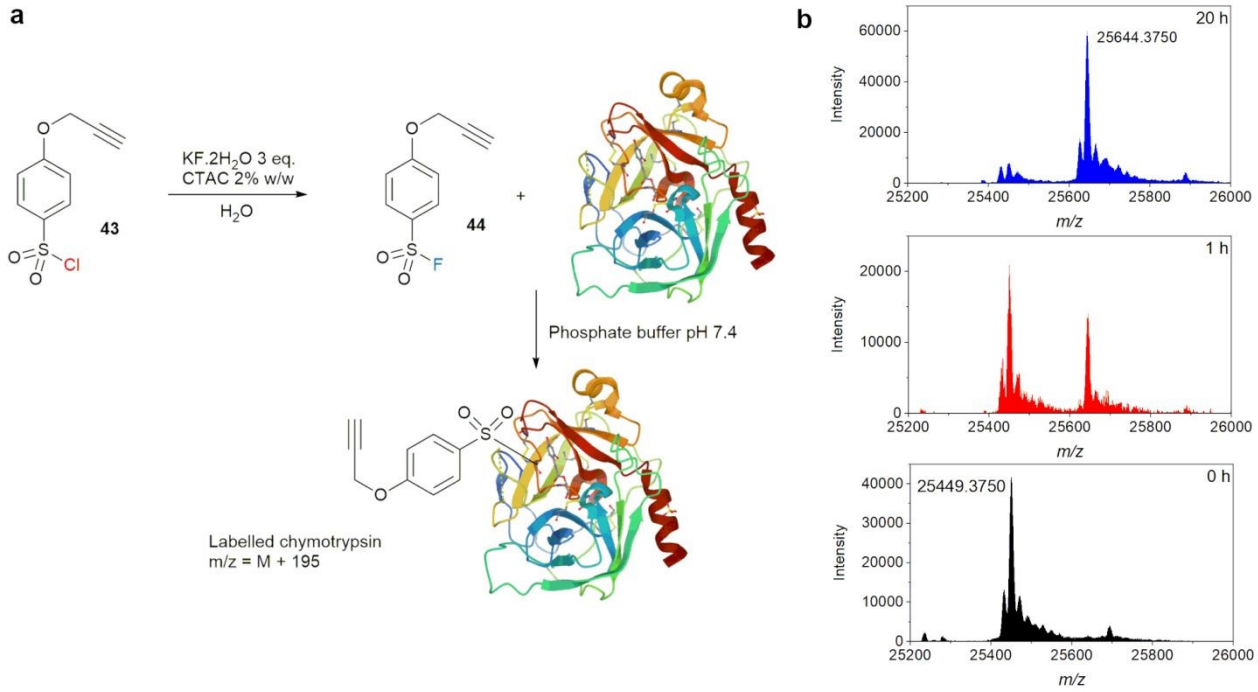


Figure 8. (a) Labelling experiments of chymotrypsin with sulfonyl fluoride generated by nucleophilic fluorination of sulfonyl chloride in water, (b) MS evolution of labelling experiment.

Author Contributions

BNN, AJB, NK and IA provided direction and supervision in the project. MHW and SLW planned and supervised the labelling experiments of chymotrypsin. AMC and KS planned and carried out

the experiments. SB carried out PM6 calculations. JM performed microscope measurements. AMC and KS contributed equally to this work.

Conflicts of interest

The authors declare the following conflict of interest: IA is an employee of AstraZeneca and may hold share options and shares in AstraZeneca.

Acknowledgements

The authors thank the EPSRC for grant support (EP/S013768/1). AMC thanks the EPSRC and AstraZeneca for a CASE studentship. We thank Dr Nicole Hondow (Univer-sity of Leeds) for her help with TEM and SEM measurements. We thank Dr Christopher Pask and Ms Jennie Dickinson for their help with DOSY experiments. Special thanks go to Prof. Brent Murray and Prof. Steven Abbott (University of Leeds) for their insightful discussions and comments into the nature of the emulsions during reaction. We also thank Croda for their generous donation of surfactants.

Notes and references

‡ Footnotes relating to the main text should appear here. These might include comments relevant not central to the matter under discussion, limited experimental and spectral data, and crystallographic data.

§

§§

etc.

1 E. J. Fendler and J. H. Fendler, ed. V. B. T.-A. in P. O. C. Gold, Academic Press, 1970, vol. 8, pp. 271–406.

2 J. H. Fendler and E. J. Fendler, eds. J. H. Fendler and E. J. B. T.-C. in M. and M. S. Fendler, Academic Press, 1975, pp. 86–103.

3 G. La Sorella, G. Strukul and A. Scarso, *Green Chem.*, 2015, **17**, 644–683.

4 F. Gallou, N. A. Isley, A. Ganic, U. Onken and M. Parmentier, *Green Chem.*, 2016, **18**, 14–19.

5 F. Gallou, P. Guo, M. Parmentier and J. Zhou, *Org. Process Res. Dev.*, 2016, **20**, 1388–1391.

6 C. M. Gabriel, N. R. Lee, F. Bigorne, P. Klumphu, M. Parmentier, F. Gallou and B. H. Lipshutz, *Org. Lett.*, 2017, **19**, 194–197.

7 B. H. Lipshutz, G. T. Aguinaldo, S. Ghorai and K. Voigtritter, *Org. Lett.*, 2008, **10**, 1325–1328.

8 P. Klumphu and B. H. Lipshutz, *J. Org. Chem.*, 2014, **79**, 888–900.

9 B. H. Lipshutz, S. Ghorai, A. R. Abela, R. Moser, T. Nishikata, C. Duplais, A. Krasovskiy, R. D. Gaston and R. C. Gadwood, *J. Org. Chem.*, 2011, **76**, 4379–4391.

10 C. Duplais, A. Krasovskiy and B. H. Lipshutz, *Organometallics*, 2011, **30**, 6090–6097.

11 C. M. Gabriel, M. Keener, F. Gallou and B. H. Lipshutz, *Org. Lett.*, 2015, **17**, 3968–3971.

12 C. Ceriani, E. Ghiglietti, M. Sassi, S. Mattiello and L. Beverina, *Org. Process Res. Dev.*, 2020, **24**, 2604–2610.

13 E. B. Tada, N. Ouarti, P. L. Silva, I. B. Blagoeva, O. A. El Seoud and M.-F. Ruasse, *Langmuir*, 2003, **19**, 10666–10672.

14 N. J. Buurma, A. Martin Herranz and J. B. F. N. Engberts, *J. Chem. Soc. Perkin Trans. 2*, 1999, 113–120.

15 S. Possidonio, F. Siviero and O. A. El Seoud, *J. Phys. Org. Chem.*, 1999, **12**, 325–332.

16 O. A. El Seoud, M.-F. Ruasse and S. Possidonio, *J. Phys. Org. Chem.*, 2001, **14**, 526–532.

17 H. Peacock and S. A. Blum, *J. Am. Chem. Soc.*, 2022, **144**, 3285–3296.

18 C.-G. Zhan and D. A. Dixon, *J. Phys. Chem. A*, 2004, **108**, 2020–2029.

19 M. T. Morales-Colón, Y. Y. See, S. J. Lee, P. J. H. Scott, D. C. Bland and M. S. Sanford, *Org. Lett.*, 2021, **23**, 4493–4498.

20 A. Talko and M. Barbasiewicz, *ACS Sustain. Chem. Eng.*, 2018, **6**, 6693–6701.

21 Q. Zheng, J. Dong and K. B. Sharpless, *J. Org. Chem.*, 2016, **81**, 11360–11362.

22 RDKit, 2023, Open-Source Cheminformatics.

23 Swain Matt, Python interface for the Chemical Identifier Resolver, https://cirpy.readthedocs.io/en/latest/.

24 Pymol.org.

25 B. M. Soares, A. M. Aguilar, E. R. Silva, M. D. Coutinho-Neto, I. W. Hamley, M. Reza, J. Ruokolainen and W. A. Alves, *Phys. Chem. Chem. Phys.*, 2017, **19**, 1181–1189.

26 K. Manabe, S. Iimura, X.-M. Sun and S. Kobayashi, *J. Am. Chem. Soc.*, 2002, **124**, 11971–11978.

27 C. Varszegi, M. Ernst, F. van Laar, B. F. Sels, E. Schwab and D. E. De Vos, *Angew. Chemie-International Ed.*, 2008, **47**, 1477–1480.

28 L. J. Diorazio, D. R. J. Hose and N. K. Adlington, *Org. Process Res. Dev.*, 2016, **20**, 760–773.

29 N. Fey, A. C. Tsipis, S. E. Harris, J. N. Harvey, A. G. Orpen and R. A. Mansson, *Chem. – A Eur. J.*, 2006, **12**, 291–302.

30 N. Fey, M. Garland, J. P. Hopewell, C. L. McMullin, S. Mastroianni, A. G. Orpen and P. G. Pringle, *Angew. Chemie Int. Ed.*, 2012, **51**, 118–122.

31 N. Fey, A. Koumi, A. V Malkov, J. D. Moseley, B. N. Nguyen, S. N. G. Tyler and C. E. Willans, *Dalt. Trans.*, 2020, **49**, 8169–8178.

32 J. R. A. Kincaid, M. J. Wong, N. Akporji, F. Gallou, D. M. Fialho and B. H. Lipshutz, *J. Am. Chem. Soc.*, 2023, **145**, 4266–4278.

33 Y. Zhang, B. Zhu, Y. Zheng and S. Huang, *J. Organomet. Chem.*, 2022, **965–966**, 122321.

34 A. Ranaudo, C. Greco, G. Moro, A. Zucchi, S. Mattiello, L. Beverina and U. Cosentino, *J. Phys. Chem. B*, 2022, **126**, 9408–9416.

35 A. Cavarzan, A. Scarso, P. Sgarbossa, R. A. Michelin and G. Strukul, *ChemCatChem*, 2010, **2**, 1296–1302.

36 W. C. Griffin, *J. Soc. Cosmet. Chem.*, 1949, **1**, 311–326.

37 J. T. Davies, in *Proceedings of the International Congress of Surface Activity*, 1957, pp. 426–438.

38 F. L. Hirshfeld, *Theor. Chim. Acta*, 1977, **44**, 129–138.

39 M. E. Tipping and C. M. Bishop, *J. R. Stat. Soc. Ser. B (Statistical Methodol.*, 1999, **61**, 611–622.
40 S. Oba, M. Sato, I. Takemasa, M. Monden, K. Matsubara and S. Ishii, *Bioinformatics*, 2003, **19**, 2088–2096.
41 M. Scholz, F. Kaplan, C. L. Guy, J. Kopka and J. Selbig, *Bioinformatics*, 2005, **21**, 3887–3895.
42 H. Wold, *Multivar. Anal.*, 1966, 391–420.
43 M. Scholz, Robustness & reproducibility, https://faq.nlpca.org/faq/robustness-reproducibility.
44 B. Ališ, J. Petrovčič, J. Jelen, G. Tavčar and J. Iskra, *J. Org. Chem.*, 2022, **87**, 5987–5993.
45 J.-W. Lee, M. T. Oliveira, H. Bin Jang, S. Lee, D. Y. Chi, D. W. Kim and C. E. Song, *Chem. Soc. Rev.*, 2016, **45**, 4638–4650.
46 G. Roagna, D. M. H. Ascough, F. Ibba, A. C. Vicini, A. Fontana, K. E. Christensen, A. Peschiulli, D. Oehlrich, A. Misale, A. A. Trabanco, R. S. Paton, G. Pupo and V. Gouverneur, *J. Am. Chem. Soc.*, 2020, **142**, 14045–14051.
47 G. Pupo, F. Ibba, D. M. H. Ascough, A. C. Vicini, P. Ricci, K. E. Christensen, L. Pfeifer, J. R. Morphy, J. M. Brown, R. S. Paton and V. Gouverneur, *Science (80-.).*, 2018, **360**, 638–642.
48 S. Kirschner, M. Peters, K. Yuan, M. Uzelac and M. J. Ingleson, *Chem. Sci.*, 2022, **13**, 2661–2668.
49 J. Dong, L. Krasnova, M. G. Finn and K. B. Sharpless, *Angew. Chemie Int. Ed.*, 2014, **53**, 9430–9448.
50 D. C. Owsley, G. K. Helmkamp and S. N. Spurlock, *J. Am. Chem. Soc.*, 1969, **91**, 3606–3609.
51 T. S.-B. Lou and M. C. Willis, *Nat. Rev. Chem.*, 2022, **6**, 146–162.
52 A. Narayanan and L. H. Jones, *Chem. Sci.*, 2015, **6**, 2650–2659.
53 D. Landini, A. Maia, F. Montanari and F. M. Pirisi, *J. Chem. Soc. Perkin Trans. 2*, 1980, 46–51.
54 M. E. Aberlin and C. A. Bunton, *J. Org. Chem.*, 1970, **35**, 1825–1828.
55 A. M. Gold and D. Fahrney, *Biochemistry*, 1964, **3**, 783–791.
56 A. J. Brouwer, T. Ceylan, A. M. Jonker, T. van der Linden and R. M. J. Liskamp, *Bioorg. Med. Chem.*, 2011, **19**, 2397–2406.
57 M. Kheirabadi, G. S. Creech, J. X. Qiao, D. S. Nirschl, D. K. Leahy, K. M. Boy, P. H. Carter and M. D. Eastgate, *J. Org. Chem.*, 2018, **83**, 4323–4335.