High fidelity Suzuki-Miyaura coupling for the synthesis of DNA encoded libraries enabled by micelle forming surfactants.

James H. Hunter,† Lisa Prendergast,‡ Louis F. Valente,ψ Andrew Madin,§ Garry Pairaudeau§ and Michael J. Waring†*

†Cancer Research UK Drug Discovery Unit, Newcastle University Centre for Cancer, Chemistry, School of Natural and Environmental Sciences, Bedson Building, Newcastle University, Newcastle upon Tyne, NE1 7RU, UK.
‡Cancer Research UK Drug Discovery Unit, Newcastle University Centre for Cancer, Newcastle University, Paul O’Gorman Building, Framlington Place, Newcastle upon Tyne, NE2 4AD, UK
ψSAS Institute Inc, JMP Division, 100 SAS Campus Drive, Cary, NC 27513, USA
§Discovery Sciences IMED Biotech Unit, AstraZeneca, 310 Cambridge Science Park, Milton Road, Cambridge, CB4 0WG, UK

KEYWORDS DNA-encoded libraries, Suzuki-Miyaura coupling, micellar mediated processes.

ABSTRACT: DNA encoded chemical libraries provide a highly efficient means of screening vast numbers of small molecules against an immobilised protein target. Their potential is currently restricted by the constraints of carrying out library synthesis in the presence of attached DNA tags, for which a limited number of reactions and substrates can be used. Even established reactions, such as Suzuki-Miyaura couplings, do not give efficient coupling reactions across a wide range of substrates and can lead to significant DNA degradation. We developed an efficient protocol for carrying out Suzuki-Miyaura couplings on DNA tagged substrates that proceeds with unprecedented efficiency to the desired biaryl products (>98% on average with no detectable DNA degradation) across a wide range of drug-like substrates using a micellar promoted process with commercial TPGS-750-M surfactant. We have demonstrated the applicability of this method in DEL synthesis by preparing a prototypical 2-dimensional 36-membered library employing the Suzuki-Miyaura coupling methodology as the final library synthesis step. This work shows, for the first time, that standard micellar surfactants enable reactions for encoded library synthesis, leading to libraries of exceptional fidelity, and demonstrates the potential to expand the range of accessible DNA compatible chemistry.

DNA encoded chemical libraries (DELs) are becoming established as an effective method of finding chemical ligands for proteins, with great potential in drug discovery and chemical biology applications.1,2,3,4,5 In this paradigm, large (potentially 106 to 109 membered) libraries are synthesised using sequential split and pool combinatorial synthetic operations that add building blocks to molecules containing a DNA tag. At each step, a DNA sequence is ligated to the DNA tag, which is unique to the building block being added, such that the overall DNA sequence at the end of the library synthesis codes for the individual structures of the appended small molecules (Figure 1a). The resulting library of compounds can then be pooled and screened as a mixture against an immobilised protein by affinity selection. Binding ligands are identified post-selection by PCR amplification and DNA sequencing. Hence, vast numbers of compounds can be screened efficiently using very small quantities of compounds and protein.

The effectiveness of this approach depends critically on the quality of the screening libraries, which is currently limited by the synthetic operations that can be carried out in the presence of the DNA tag.6,7 The library synthesis must be carried out in aqueous media at high dilution and must not contain reagents that are capable of reacting with DNA. For the affinity selection to yield high fidelity information, the reactions must proceed to full conversion across a wide range of substrates such that each library member is present in similar concentration in the final library. Recent studies have shown that current, commonly employed methodologies for DEL synthesis are far from ideal in this regard and can result in significant loss of the amount of amplifiable DNA.6 The scope and potential of DELs could be expanded significantly with the establishment of more general, versatile synthetic methods that are compatible with DEL conditions and proceed to full conversion across a wide range of substrates.
Micellar catalysis is increasingly being established as a means of carrying out organic reactions in aqueous media. This has potential advantages over reactions in organic solvents; efficient reactions can be brought about by the localisation of organic reactants inside the micelles, resulting in high effective concentrations in a poorly solvated environment. We hypothesised that this protocol could be used in DEL synthesis, widening the scope of reactants that could be employed by using the micelles to dissolve non-

Figure 1. a) Split and pool DEL-synthesis scheme; b) Model of the micelle enabled process

Figure 2. a) Reaction scheme for the test DNA-linked substrate iodobenzene 1 and formation of desired biaryl 2 and side product 3; b) Chromatograms of couplings of 1 with phenyl boronic acid a to give product 2a under literature and Nok micellar conditions.

Accordingly, we selected the Suzuki-Miyaura coupling as an ideal test case for the application of micellar surfactant mediated reactions to DEL synthesis. The reaction has been reported to proceed well under micellar conditions on standard organic substrates. An aryl iodide substrate, linked to a short double-stranded DNA sequence that has been previously applied in DELs, was prepared (Figure 2a and Supporting Information). Coupling of the resulting aryl iodide with phenyl boronic acid under the literature conditions failed to yield the desired product 2a, indicating that the reaction was suboptimal.

A related concept has recently been established in which Bronsted acids coupled covalently to block copolymer surfactants have been used for DEL synthesis. This protocol requires the bespoke synthesis of reagents immobilised on the surfactants. We believed that this may be unnecessary in some cases, and that unmodified, commercial surfactants could mediate a similar effect by localisation of reagents through non-covalent interactions, leading to a process that would be operationally far simpler and more generally applicable.

The Suzuki-Miyaura cross-coupling reaction has been shown to be a reliable and versatile reaction in traditional synthetic medicinal chemistry and has been employed for the synthesis of DELs. This work shows that couplings are suboptimal. Many substrates do not proceed with complete conversion, resulting in the starting aryl halides being retained in the product mixtures. This is clearly problematic in multistep library synthesis, in which the starting materials cannot easily be removed. In two reports, coupling aryl iodides to a range of boronic acids resulted in >90% conversion to desired product in only 37% or 66% of cases, with heterocyclic substrates, perhaps most desirable as library components, being particularly problematic. Moreover, these conversions neglect the formation of by-products such as dehalogenation and degradation of DNA, which can be significant. It has been shown that Suzuki-Miyaura coupling conditions can destroy 70% of amplifiable DNA.

Accordingly, we selected the Suzuki-Miyaura coupling as an ideal test case for the application of micellar surfactant mediated reactions to DEL synthesis. The reaction has been reported to proceed well under micellar conditions on standard organic substrates. An aryl iodide substrate, linked to a short double-stranded DNA sequence that has been previously applied in DELs, 1 was prepared (Figure 2a and Supporting Information). Coupling of the resulting aryl iodide with phenyl boronic acid under the literature conditions failed to yield the desired product 2a, indicating that the reaction was suboptimal.
conditions resulted in formation of only 26% of the desired biphenyl 2a (Ar = Ph) with 39% starting material 1 remaining. Deiodinated 3 was also observed (34%) along with further uncharacterised impurities (Figure 2b). In contrast, carrying out the coupling in the presence of Nok micelles resulted in complete conversion to biphenyl 2a with no detectable impurities and with the complementary DNA strand remaining fully intact.

Encouraged by this result, we explored the coupling in the presence of Nok across a wider range of boronic acids (Table S1). The conditions did not prove to be general, the reaction with substituted phenyl or a range of heterocyclic boronic acids showed remaining starting material 1 or dehalogenated 3 to significant degrees. We therefore explored the reaction conditions further using the poorly performing 4-trifluoromethylphenyl boronic acid b. Varying the base had significant effects, with inorganic bases generally leading to improved levels of conversion (K3PO4 72% product, LiOH 88%, Table S2). Changing the micellar surfactant from Nok to TPGS-750-M resulted in a moderate improvement in conversion with potassium phosphate but not the lithium hydroxide (Table S3). Addition of 15% THF, reported to induce swelling of the micelles, with increased boronic acid and potassium phosphate resulted in 100% product 2b formation without detectable dehalogenation.

Having developed conditions that produced 100% conversion with 4-trifluoromethylphenyl boronic acid b, an extensive range of substrates was profiled (Table 1). This revealed that the reaction worked well for phenyl boronic acids, resulting in 100% conversion to the biaryl product for phenyl a, 4-methoxyphenyl c, 4-chlorophenyl d, 4-fluorophenyl e, 3-trifluoromethylphenyl g boronic acids. 2-methylphenyl boronic acid h also gave good (97%) conversion. 4-Carboxyphenyl boronic acid f gave less efficient coupling. Heterocyclic boronic acids also performed less well, indol-5-yl boronic acid i gave 100% conversion but electron poor pyrimidinyl j and pyridyl k boronic acids gave poor conversion with significant dehalogenation. Phenyl boronic acid pinacol ester l gave 100% conversion but the pyrazolyl boronic acid pinacol ester m gave poor conversion with significant remaining starting material. MIDA boronates, which are reported to perform well for couplings of heterocyclic systems prone to protodeboronation, performed poorly for both phenyl n and heteroaryl o and p systems, resulting in low conversion.

Whilst these results were promising, for the synthesis of lead-like libraries, it would clearly be far more desirable to develop conditions that also gave efficient conversion for heteroaryl substrates. Accordingly, we sought to further optimise the reaction conditions. Reducing the amount of base to 312 equivalents, in an attempt to suppress putative boronate hydrolysis, resulted in 100% conversion for 5-pyrimidinyl boronic acid j but gave less efficient reactions for all the other substrates (Table S4). Increasing the amount of palladium to 15 equivalents gave 100% conversion to product for pyridyl k but only 60% for pyrimidinyl j (Table S5).

Table 1. Coupling of 1 with a range of boronic acids, conditions: Pd(dtbpf)Cl2 (2 eq.), boronate (3125 eq.), K3PO4, (4688 eq.), 2% TPGPS-750-M, 15% THF, 50 °C.

<table>
<thead>
<tr>
<th>Boronate</th>
<th>% product 2</th>
<th>% 1</th>
<th>% 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>B(OH)2</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>b</td>
<td>B(OH)2</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>c</td>
<td>B(OH)2</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>d</td>
<td>B(OH)2</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>e</td>
<td>B(OH)2</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>f</td>
<td>B(OH)2</td>
<td>69*</td>
<td>0</td>
</tr>
<tr>
<td>g</td>
<td>B(OH)2</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>h</td>
<td>B(OH)2</td>
<td>97</td>
<td>0</td>
</tr>
<tr>
<td>i</td>
<td>B(OH)2</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>j</td>
<td>B(OH)2</td>
<td>34</td>
<td>0</td>
</tr>
<tr>
<td>k</td>
<td>B(OH)2</td>
<td>79</td>
<td>0</td>
</tr>
<tr>
<td>l</td>
<td>B(OH)2</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>m</td>
<td>B(OH)2</td>
<td>5</td>
<td>95</td>
</tr>
<tr>
<td>n</td>
<td>B(OH)2</td>
<td>65</td>
<td>0</td>
</tr>
<tr>
<td>o</td>
<td>B(OH)2</td>
<td>6</td>
<td>94</td>
</tr>
<tr>
<td>p</td>
<td>B(OH)2</td>
<td>67</td>
<td>0</td>
</tr>
</tbody>
</table>

*Formation of an unknown by-product of 102.8 Da higher mass than 2f and an analogous adduct on the complementary DNA strand.
We applied the optimised conditions from the FED to the full range of boronic acid substrates (Table 2). Gratifyingly, this led to exceptionally efficient reaction across the full range of substrates, resulting in >95% desired product in all cases and, in 12 out of 18 examples, the desired product was the only one detectable. This included electron poor (b, d, e, f, g), electron-rich (c, h), ortho-substituted (h), acidic (f) and heterocyclic (l, j, k, m, o, p, q, r) substrates and applied equally well to pinacolato- (l, m) and MIDA- (o, p, q, r) boronate esters. The only exception to this was 2-chloropyridin-3-yl MIDA boronate p, which demonstrated further coupling of a second boronate with the chloro-substituent. The results with pyrazol-3-yl pinclatoborate m and 4-methylpyridin-3-yl MIDA boronate o are particularly noteworthy as they had previously been the least reactive substrates (<10% conversion) and afforded >95% product under the optimised conditions.

To establish that the surfactant was still important for this outcome, we repeated the reactions using identical conditions except for leaving out either the surfactant, THF co-solvent or both. In all cases the outcomes were inferior across the range of substrates (Figure 4 and Table S7). In some cases, efficient reactions could still be observed with THF or surfactant alone (boronates a-l gave >90% conversion in one or both cases). In other cases, reactions were significantly less efficient, resulting in appreciable amounts of starting material 1 and dehalogenated 3 (m-r). Overall, the reactions in the absence of surfactant gave on average 89% product 2, 4% starting material 1 and 6% dehalogenated 3; leaving out the THF gave 91% 2, 2% 1 and 5% 3 and leaving both out resulted in 63% 2, 36% 1 and 2% 3. In comparison, the optimised micellar conditions led to >98% product on average with ≤0.5% of any other product (Figure 4b).

Table 2. Coupling of 1 with a range of boronic acids, conditions: Pd(dtbpf)Cl2 (11 eq.), boronate (750 eq.), K2PO4, (800 eq.), 2% TPGPS-750-M, 15% THF, 60 °C.

<table>
<thead>
<tr>
<th>Boronate</th>
<th>% product 2</th>
<th>% 1</th>
<th>% 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>97</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>b</td>
<td>100</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Figure 3. Optimisation of coupling conditions leading to desired biaryl product 2j and 2k for heterocyclic boronates using FED. a) Cube plots showing the percentage of desired product for 5-pyrimidinyl and 3-pyridyl boronic acids j and k showing the degree of conversion across the design space for the modelled data; b) Comparison of fitted data from the predictive model with the experimental data; c) Assessment of the most influential terms in the FED and profile of the effect of temperature, palladium equivalents and base equivalents for 3-pyridyl boronic acid k; d) Response curves showing relationship between % product formation and individual parameters for pyridyl boronic acid and robustness of response between j and k for optimized conditions (Temp 60 °C, 11 eq. Pd and 800 eq. base); e) Contour plots showing the combined effect of temperature and palladium equivalents for each substrate.

With the evidence of co-dependencies between reaction conditions and substrate, we elected to further optimise the conditions systematically using factorial experimental design (FED). A response surface design exploring the variables of temperature (40 to 60 °C), base equivalents (100 to 1500) and palladium equivalents (2 to 20) along with varying the heterocyclic boronate (5-pyrimidinyl j or 3-pyridyl k) was conducted (Figure 3 and Table S6). This revealed clear differences between optimal conditions for each boronic acid. 3-Pyridyl boronic acid k favoured higher palladium and base equivalents, whereas 5-pyrimidinyl j gave higher conversions with lower palladium and base. In all cases the relationships were parabolic, with optima within the design range. Both favoured higher temperature. The experiment indicated that the best conditions for the efficient coupling of both boronic acids was close to the centre point for palladium and base at higher temperature (1 eq. palladium, 800 eq. base at 60 °C). It also indicated second order effects between the palladium equivalents / temperature, base equivalents / boronic acid and palladium equivalents / boronic acid.
c $\text{MeO}$ 100 0 0

d $\text{Cl}$ 100 0 0

e $\text{F}$ 100 0 0

f $\text{HO}_2\text{C}$ 100 0 0

g $\text{F}_3\text{C}$ 100 0 0

h 100 0 0

i 100 0 0

j 100 0 0

k 100 0 0

l 96 0 0

m 97 3 0

n $\text{HN}$ 96 0 4

o 100 0 0

p $\text{Cl}$ 81* 0 1

q 98 0 2

r 100 0 0

* 18% further coupling of boronate with the chloro-functionalitity of the product.

Figure 4. Confirmation of the promotion of reaction by TPGS, a) % formation of 2a-r under optimised micellar conditions (red) compared to the analogous reaction in the absence of THF (blue), TPGS (green) or both (amber) across the selected substrates, b) average % formation of product 2a-r (red) relative to starting material 1 (green), dehalogenated 3 (blue) or other side products (amber) for the optimised micellar conditions compared to the analogous reaction in the absence of THF, TPGS or both.

The optimised conditions were used to develop a prototypical DEL. This consisted of a 2-dimensional, 36-member library design employing addition of 6 arylhalide-bearing carboxylic acids (s, t, u, v, w, x) to an amino-tagged DNA headpiece followed by Suzuki-Miyaura coupling of 6 aryl boronates (a, b, c, j, m, q, Figure 5). At both stages, a coding sequence unique to each building block (Tables S9 and S10) was ligated followed by the addition of a closing sequence resulting in a library of 36 compounds each appended to a unique DNA tag.

The fidelity of the DNA tags upon completion of the library synthesis was confirmed by PCR amplification. Analysis of the amplified DNA by gel electrophoresis showed the only significant band to be ca. 140 base pairs, corresponding to the fully intact final DNA sequences (Figure S1). The amplified DNA was sequenced by next generation sequencing, which identified the coding amplicons for all 36 members of the library (Table S11). This confirms that the DNA remains amplifiable and readable after subjection to the library synthesis conditions, including the optimised Suzuki-Miyaura reaction. The read frequency for both the conjugated strand and the complementary strand were comparable, proving that the linker attachment does not affect the amplification of the DNA. Together this establishes that this protocol amenable to preparation of screening libraries.

These results provide a paradigm shift in the efficiency of on-DNA Suzuki-Miyaura couplings for the synthesis of DELs. Our conditions result in exceptionally high levels of product formation with no detectable DNA damage (by analysis of the product mass spectra). This will enable the preparation of libraries of far greater fidelity in the future, which we anticipate will lead to much more reliable biological screens. The results provide the first indication that commercial micellar surfactants can be of general utility.
for DEL synthesis and may allow the expansion of the current range of DEL compatible chemistry as well as improving the scope and efficiency of existing reactions.

Figure 5. Synthesis of a prototypical 36 member DEL via amide coupling and micellar promoted Suzuki Miyaura reaction. Conditions: (i) DNA ligation of acid codon (green), then carboxylic acid, EDC, HOAt, DIPEA, MOPS, DMSO water; (ii) DNA ligation of boronate codon (blue) and closing sequence, then Pd(dtbpf)Cl$_2$ (11 eq.), boronate (750 eq.), K$_3$PO$_4$, (800 eq.), 2% TPGS-750-M, 15% THF, 60 °C.

ASSOCIATED CONTENT

Experimental for optimised coupling procedure
An aliquot of boronic acid solution (20 µl, 0.75 M in DMF) was added to a 50 µl glass insert for a para-dox 96-well micro Para-dox™ photoredox/optimisation plate. The DMF was then removed at 55°C in a centrifugal evaporator for 30 mins. To this solution was added 5% TPGS-750-M in water (4 µl), potassium phosphate (6 µl of a solution of 113.23 mg in 200 µl water) and 1 (20 µl of 1 mM in 2% TPGS-750-M in water). The vials were vortexed for 30 seconds each to enhance mixing. Pd(dtbpf)Cl$_2$ (4.5 µl of 6.37mg in 200 µl THF) was added and the samples vortexed again for 10 seconds each. The samples were then heated in a Para-dox™ 96-well micro photoredox/optimisation plate at 60°C for 5 hours. QTOF mass spectrometry was used to analyse reactions. Samples prepared by adding reaction mixture (5 µl) to water (595 µl) and filtered through a hydrophilic PTFE filter. Ethyl acetate (40 µl) was added to each and the vial vortexed. The organics were removed, aqueous sodium chloride (4 µl, 4M) and ethanol (70 µl) were added and the mixture incubated at -78°C for 1 hour. The mixture was then centrifuged and the ethanol layer removed. Ethanol (70 µl) was added and the process repeated. The pellet of DNA was then dissolved in water to give a mM solution.


Supporting Information. Detailed experimental procedures, further experimental results for the optimisation and full analytical data are provided. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

* Prof. Michael J. Waring,
Cancer Research UK Drug Discovery Unit, Newcastle University Centre for Cancer, Chemistry, School of Natural and Environmental Sciences, Bedson Building, Newcastle University, Newcastle upon Tyne, NE1 7RU, UK. Tel: +44 (0) 191 208 8591, Email: mike.waring@ncl.ac.uk

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

Funding Sources
The research was supported by AstraZeneca (PhD studentship award to JHH) and Cancer Research UK (Newcastle Drug Discovery Unit Programme Grant funding, grant reference C2115/A21421).

ACKNOWLEDGMENT
This manuscript is dedicated in memoriam to Dr. David R. Waring (1944 - 2019), a gifted chemist and inspirational father.

We thank Prof. Martin E. M. Noble, Prof. Ian Hickson, Dr. Richard Heath and Dr. Mathew Martin for advice and assistance with the DNA ligations. We thank Dr. Sirintra Nakjang for assistance with analysis of NGS data.

ABBREVIATIONS
DEL, DNA encoded library; DNA, deoxyribonucleic acid; PCR, polymerase chain reaction

REFERENCES
SYNOPSIS TOC: A highly efficient method of carrying out Suzuki-Miyaura couplings on DNA-conjugated substrates mediated by micellar surfactants was developed. The optimised protocol gives exceptionally clean (>98% on average) conversion to the desired biaryl products across a diverse range of substrates and with no detectable DNA degradation. This approach represents a step change from existing methods and will enable the production of DNA-encoded libraries with exceptionally high fidelity.

Table of Contents (TOC) graphic.

>98% product
Aryls, heteroaryls, boronic acids, pinacolato and MIDA boronate esters