ChemComm

COMMUNICATION

View Article Online
View Journal | View Issue



Cite this: *Chem. Commun.*, 2014, **50**, 5288

Received 14th October 2013,
Accepted 17th December 2013

DOI: 10.1039/c3cc47894g

www.rsc.org/chemcomm

# Synthesis of squaraine dyes under mild conditions: applications for labelling and sensing of biomolecules†

Mazen Haj Sleiman and Sylvain Ladame*

**We report the synthesis of squaraine dyes under mild conditions by carbodiimide activation of squaric acid or semi-squaraine dyes. Despite low yields when the reaction was carried out in solution, these conditions were successfully applied to efficient peptide labelling on resin and nucleic acid sensing in solution.**

Cyanine dyes had already been widely used for a broad range of applications (*e.g.* in photographic emulsion)[1] before they became increasingly popular in the 1990s as fluorescent labels for biomolecules, in particular nucleic acids.[2] They generically consist of a conjugated system based on a polymethine chain linking two nitrogen-containing heterocycles (*e.g.* indoles and benzothiazoles).[3] Squarilium dyes, first discovered by Treibs *et al.* in 1965,[4] differ from traditional cyanine dyes by the introduction of an oxocyclobutenolate ring within the polymethine chain linking both terminal heterocycles. This central, electron-deficient, cyclobutadione bridge confers extra rigidity and planarity on squaraine dyes. These zwitterionic dyes typically absorb in the red to near infra-red (630–650 nm) and are characterised by large extinction coefficients ($\geq 10^5$ $M^{-1}$ $cm^{-1}$), good photo-stability, high fluorescent quantum yields (up to 65%) and long fluorescence lifetime, especially when bound to biomacromolecules.[5] However, synthesis of these squaraine dyes typically requires harsh conditions (*e.g.* high temperatures), limiting significantly their use in biological applications. Herein, we demonstrate that indole squaraine dyes can be synthesised under mild (and aqueous) conditions using carbodiimide-activated squaric acid. We also investigated the benefits of these milder conditions for the labelling and sensing of biomolecules.

Symmetrical squarylium dyes are generally prepared by reacting one equivalent of squaric acid with two equivalents of the electron-donating heterocycle in refluxing *n*-butanol.[6] Yields can be increased by using azeotropic co-solvents (*e.g.* isopropanol + benzene or toluene) and/or by the addition of tributyl orthoformate as a drying agent.[7] The reactivity of heterocycles belonging to the indole, benzothiazole or the quinoline type is also strongly affected by the acidity of the 2-methyl hydrogen atoms. Depending on the nature of the heterocycle (and of its substitution pattern), a limited formation of the 2-methylene species leads to poor yields, due to the absence of reaction or due to the reaction stopping at the semi-squaraine (reaction of only one molecule of heterocycle per molecule of squaric acid). Attempts to further improve the yield of the reaction of squaraine dye synthesis include the use of a large excess of quinoline (to favour the formation of the reactive 2-methylene heterocyclic species), or the use of an 'activated' squaric acid: squaric esters (most commonly butyl and ethyl esters) or 3,4-dichlorocyclobut-3-ene-1,2-dione.[8] However, even under such 'optimised' conditions, reaction of dye formation only proceeds at high temperatures ($\geq 80$ °C), and preferentially under anhydrous conditions, thereby preventing its broader use for labelling and sensing of biomolecules.

Herein, we have investigated the potential of various carboxylic acid activating agents commonly used in peptide chemistry to facilitate the reaction between 1,3,3-trimethyl-2-methyleneindoline and squaric acid, chosen as a model system (Fig. 1). Briefly, a pre-activated mixture of 1 equivalent of squaric acid and two equivalents of activating agents was added to a solution of 2 equivalents of 1,3,3-trimethyl-2-methyleneindoline in various organic solvents (DCM, DMF, DMSO, and dioxane) and the reaction was monitored by HPLC. Analytically pure samples of the symmetrical squaraine dye and of the semi-squaraine dye were synthesised following procedures from the literature[8,9] and were subsequently used to determine the reaction yield. For a couple of examples, the product of the reaction was purified by flash-chromatography on silica gel and provided a yield similar to that estimated by HPLC, thus validating our HPLC-based estimations (see ESI† for detailed experimental protocols). Of all coupling agents tested (HATU, oxyma, EDCI, DCC, DIC), only those belonging to the carbodiimide family led to the formation of detectable amounts of squaraine dye **1**. The results of this screening are summarised in Table 1.

As shown in Table 1 (entries A–L), the nature of the solvent, the reagent concentration and the temperature all influenced the efficiency of the reaction of squaraine dye formation. Under comparable

*Department of Bioengineering, Imperial College London, London, SW7 2AZ, UK. E-mail: sladame@imperial.ac.uk*

† Electronic supplementary information (ESI) available: Detailed experimental procedures and product characterisation. See DOI: 10.1039/c3cc47894g

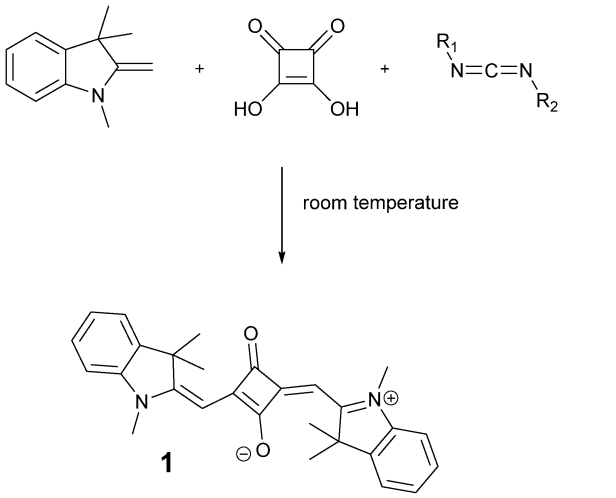


Fig. 1 General scheme for squaraine dye synthesis *via* carbodiimide-mediated activation of squaric acid.

Table 1 Effect of selected activating agents, solvents, reactant concentration and temperature on the efficiency of the reaction of symmetrical squaraine dye formation[a]

| | Activating agent | Solvent | Temp. (°C) | Conc.[b] (mM) | Calcd yield (%) |
|---|---|---|---|---|---|
| A | DIC | DMF | RT | 0.5 | <1 |
| B | DIC | DMSO | RT | 0.5 | <1 |
| C | DIC | DCM | RT | 0.5 | 2.8 |
| D | DIC | DCM | RT | 1.0 | 9.1 (10.0[c]) |
| E | DCC | DCM | RT | 1.0 | 9.0 (9.0[c]) |
| F | EDCI | DCM | RT | 1.0 | 10.6 (9.0[c]) |
| G | EDCI | Dioxane | RT | 1.0 | 1.4 |
| H | EDCI | Dioxane : $H_2O$ | RT | 1.0 | 2.2 |
| I | EDCI | $H_2O$ | RT | 1.0 | <1 |
| J | DIC | DCM | 60 | 0.5 | 14.2 |
| K | DIC | DCM | 60 | 0.5 | 11.1[d] |
| L | DIC/NHS | DCM | 60 | 0.5 | 14.9 |

[a] Reactions were carried out in a sealed tube over 24 h using 1 molar equiv. squaric acid, 2 molar equiv. carbodiimide and 2 molar equiv. indoline, unless stated otherwise. [b] Concentration of squaric acid in the reaction mixture. [c] Purified yield. [d] Reaction stopped after 1 hour.

conditions, all three carbodiimides tested performed equally well (entries D–F) with yields ranging between 9 and 11%. Using DIC as an activating agent, DCM proved to be the best solvent (entries A–C) while doubling the concentration of all reagents (from 0.5 mM to 1.0 mM in squaric acid) increased the yield of the reaction more than 3-fold (entries C and D). Using a water-soluble carbodiimide (EDCI) enabled the reaction to be performed in aqueous solvents, a condition often required for the labelling or sensing of biomacromolecules. Although only traces of dye were observed when working in pure water (entry I), yields up to 2.2% were obtained when working in a dioxane–water (1 : 1) binary mixture (entry H). Increasing the temperature led to a significant increase in the reaction efficiency. The yield of the reaction carried out in the presence of DIC increased 5-fold (14.2% *versus* 2.8%) when the temperature was increased up to 60 °C (entries C and J). Under such conditions, significantly faster kinetics was also observed with 11.1% dye formed after only 1 hour (entry K). Addition of *N*-hydroxysuccinimide (NHS) to the reaction mixture had however little effect on the reaction efficiency (entries J and L). No significant effect was either observed upon addition of drying agent $MgSO_4$.

Despite some relatively low yields in solution, we hypothesised that our mild reaction conditions would be significantly better suited than traditional approaches described in the literature (high temperatures with additives and/or dehydrating agents) for the labelling of peptides (or other biomolecules) on solid supports. As a model system, we investigated the potential of our new synthetic strategy for the fluorescent labelling of peptides on resin. A GFK-$NH_2$ tripeptide was synthesised on an aminomethyl ChemMatrix resin using standard Fmoc chemistry and was functionalised at its N-terminus with a 2-methylene indoline (see ESI† for detailed experimental conditions).

Whilst the stepwise synthesis of the squaraine dye (*via* formation of a semi-squaraine dye on resin followed by the addition of a second molecule of 2-methylene indole) proved both slow and inefficient, an alternative route was chosen that involved reaction of the immobilised peptide–indoline with a solution of semi-squaraine dye (1 to 5 molar equiv.) pre-activated with a stoichiometric amount of DIC in DCM (Fig. 2). After 6 h at room temperature, the resin was washed thoroughly, the peptide cleaved off the resin and the mixture was analysed by HPLC and mass spectrometry (MALDI). Interestingly, only traces of unreacted peptide were detectable when using 3 equiv. of activated squaric acid and near-total conversion into the desired peptide–squaraine dye conjugate was achieved when using 5 equiv. of squaric acid (Fig. 2). Next, a similar strategy was used for labelling a Peptide Nucleic Acid (PNA) oligomer on resin. PNAs are DNA (or RNA) peptidic analogues that can hybridise to complementary oligonucleotides *via* Watson–Crick base-pairing to form stable PNA–DNA (or PNA–RNA) heteroduplexes.[10] A labelling experiment was carried out with a water soluble 5-mer PNA functionalised at its N-terminus with the same 2-methylene indole. After 24 h reaction at rt, a 70% conversion into the fluorescently labelled PNA was obtained (see ESI†), demonstrating the applicability of this chemistry to various immobilised oligomers. Finally, the same strategy was applied to the synthesis of an unsymmetrical squaraine dye on resin. 2,3,3-Trimethyl-3*H*-indole was first quarternised with 6-bromohexanoic acid and immobilised on resin. Reaction with a semi-squaraine dye in the presence of DIC afforded the desired

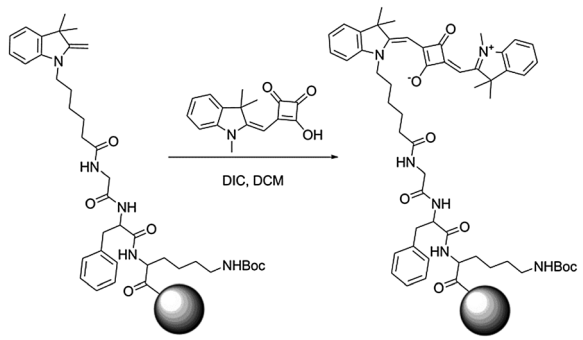


Fig. 2 Solid-phase synthesis of a peptide–squaraine dye conjugate.

unsymmetrical squaraine dye in great purity (>90%, as estimated by $^1$H NMR spectroscopy).

Having demonstrated that synthesis of squaraine dyes could proceed in aqueous solutions (see Table 1, entries H and I), although in poor yields, and considering the fluorogenic properties of the reaction, we decided to investigate its suitability for optical sensing applications based on the concept of oligonucleotide-templated reaction (OTR).[11] Widespread in Nature, the concept of OTRs has recently been applied to the highly sensitive and sequence-specific detection of nucleic acid biomarkers both *in vitro* and *in vivo*. For such applications, oligonucleotide-templated fluorogenic reactions, for which the appearance of a characteristic fluorescent signal can be directly correlated to the sensing of the unique nucleic acid target of interest, proved extremely popular.

We recently reported the use of fluorogenic PNA probes for the detection of DNA secondary structures, based on an OTR of cyanine dye (C3) formation.[12] Herein, we investigated the applicability of our reaction of squaraine dye synthesis for DNA sensing. Two water-soluble PNAs functionalised at their N or C terminus with a semi-squaraine dye (PNA1) and with a 2-methyleneindoline moiety (PNA2), respectively, were synthesised on solid phase (see ESI† for sequences and experimental details). Reactions between PNA1 and PNA 2 (70 μM each) in potassium phosphate buffer (15 mM, pH 7.4) containing an excess of EDCI (700 μM, 10 equiv.), in the absence and in the presence of a complementary DNA strand (70 μM, 1 equiv.), were monitored by fluorescence spectroscopy ($\lambda_{exc}$ = 610 nm). Whilst at such low concentration of PNA probes, no significant formation of the squaraine dye was detectable in the absence of a template; a fluorescent signal characteristic of squaraine dyes ($\lambda_{em}$ = 645 nm) was observed after 12 h. When raising the temperature up to 40 °C, an 8-fold increase in the fluorescence intensity was observed after 3 h upon addition of the DNA template (when compared to the non-templated reaction) (Fig. 3). The product of the OTR was finally characterised by MALDI, thus confirming that the observed fluorescence at $\lambda_{em}$ = 645 nm originated, as anticipated, from the formation of the PNA1–squaraine–PNA2 conjugate ($[M + H]^+$ $m/z$ = 3990, yield < 5%, as estimated by MALDI, see ESI†). Application and optimisation of this new fluorogenic OTR in the sensing of nucleic acid secondary structures is currently underway in our laboratory.

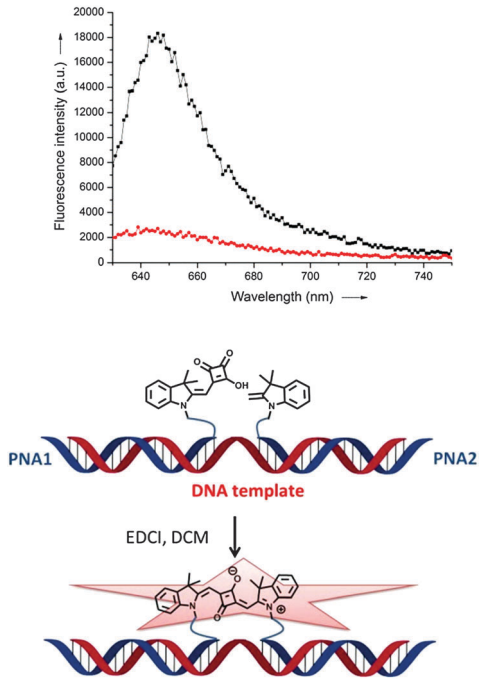


Fig. 3 (bottom) DNA-templated synthesis of a squaraine dye from two non-fluorescent precursors; and (top) fluorescence emission spectra of a stoichiometric mixture of PNA1 and PNA2 (70 μM each) in potassium phosphate buffer (15 mM, pH 7.4) in the absence (red circles) or in the presence (black squares) of a DNA template. Spectra were recorded after 3 h at 40 °C and after a 6-fold dilution in water ($\lambda_{exc}$ = 610 nm).

We have demonstrated that symmetrical squaraine dyes can be synthesised under mild (and aqueous) conditions *via* carbodiimide-mediated activation of squaric acid (or of the semi-squaraine dye). Although yields of reactions carried out in solution remain low (up to 15%), near quantitative conversions are obtained when the reaction is carried out on a solid-support. We also show that the fluorogenic reaction of squaraine dye synthesis can be applied to DNA sensing, using a strategy based on the concept of OTR. Although more work is required to further improve the sensitivity of the system, this first example of DNA-templated synthesis of a squaraine dye emitting in the NIR from two non-fluorescent precursors is likely to have valuable applications for the development of new DNA (or RNA)-targeting probes.

This research was supported by the Leverhulme Trust (grant RPG-2012-603) and by the EU FP7 (Career Integration Grant CIG 293981).

## Notes and references

1 J. A. Leermakers, B. H. Carroll and C. J. Staud, *J. Chem. Phys.*, 1937, **5**, 878.
2 For an early example see R. B. Mujumdar, L. A. Ernst, S. R. Mujumdar, C. J. Lewis and A. S. Waggoner, *Bioconjugate Chem.*, 1993, **4**, 105.
3 F. M. Hamer, in *The Cyanine Dyes and Related Compounds*, ed. A. Weissberger, Interscience, London U.K., 1964.
4 A. Treibs and K. Jacob, *Angew. Chem., Int. Ed.*, 1965, **4**, 694.
5 See recent review L. Hu, Z. Yan and H. Xu, *RSC Adv.*, 2013, **3**, 7667.
6 G.-S. Jiao, A. Loudet, H. B. Lee, S. Kalinin, L. B. A. Johansson and K. Burgess, *Tetrahedron*, 2003, **59**, 3109.
7 H.-E. Sprenger and W. Ziegenbein, *Angew. Chem., Int. Ed.*, 1967, **6**, 553; K. Y. Law and F. C. Bailey, *J. Org. Chem.*, 1992, **57**, 3278.
8 S. Yagi, Y. Hyodo, S. Matsumoto, N. Takahashi, H. Kono and H. Nakazumi, *J. Chem. Soc., Perkin Trans. 1*, 2000, 599.
9 M. Matsui, M. Fukushima, Y. Kubota, K. Funabiki and M. Shiro, *Tetrahedron*, 2012, **68**, 1931.
10 M. Egholm, O. Buchardt, L. Christensen, C. Behrens, S. M. Freier, D. A. Driver, R. H. Berg, S. K. Kim, B. Norden and P. E. Nielsen, *Nature*, 1993, **365**, 566.
11 For recent reviews see C. Percivalle, J. F. Bartolo and S. Ladame, *Org. Biomol. Chem.*, 2013, **11**, 16; K. Gorska and N. Winssinger, *Angew. Chem., Int. Ed.*, 2013, **52**, 6820. For selected recent examples see A. Shibata, T. Uzawa, Y. Nakashima, M. Ito, Y. Nakano, S. Shuto, Y. Ito and H. Abe, *J. Am. Chem. Soc.*, 2013, **135**, 14172; K. K. Sadhu and N. Winssinger, *Chem.–Eur. J.*, 2013, **19**, 8182; A. Roloff and O. Seitz, *ChemBioChem*, 2013, **14**, 2322. For selected earlier examples seeM. K. Herrlein, J. S. Nelson and R. L. Letsinger, *J. Am. Chem. Soc.*, 1995, **117**, 10151; Z. A. Shabarova, I. N. Merenkova, T. S. Oretskaya, N. I. Sokolova, E. A. Skripkin, E. V. Alexeyeva, A. G. Balakin and A. A. Bogdanov, *Nucleic Acids Res.*, 1991, **19**, 4247; S. Sando and E. T. Kool, *J. Am. Chem. Soc.*, 2002, **124**, 2096; Z. J. Gartner, M. W. Kanan and D. R. Liu, *J. Am. Chem. Soc.*, 2002, **124**, 10304.
12 K. Meguellati, G. Koripelly and S. Ladame, *Angew. Chem., Int. Ed.*, 2010, **49**, 2738; G. Koripelly, K. Meguellati and S. Ladame, *Bioconjugate Chem.*, 2010, **21**, 2103.

 