Photochemistry and Photobiology, 2023, 99: 562–569

Special Issue Research Article

Squaraine Dyes as Fluorescent Turn-on Probes for Mucins: A Step Toward Selectivity†

Cosmin Butnarasu[1], Carlotta Pontremoli[2], Maria Jesus Moran Plata[2], Nadia Barbero*[2] and Sonja Visentin[1]

[1]Department of Molecular Biotechnology and Health Science, University of Turin, Torino, Italy
[2]Department of Chemistry, NIS Interdepartmental and INSTM Reference Centre, University of Torino, Torino, Italy

Received 15 July 2022, accepted 15 September 2022, DOI: 10.1111/php.13722

ABSTRACT

Mucins are a family of long polymeric glycoproteins which can be overexpressed in several types of cancers, and over recent years, great attention was addressed to identify mucins as an important biomarker of adverse prognosis. Fluorometric detection mediated by fluorescent probes could represent a winning strategy in the early diagnosis of different pathologies. Among promising biological fluorescent probes, squaraines are gaining particular attention, thanks to their sharp and intense absorption and emission in the NIR region. In this contribution, three squaraine dyes bearing different substituents and with different lipophilicity have been investigated for their ability to detect mucin. The turn-on response upon the addition of mucin has been investigated by means of absorbance and fluorescence spectroscopy. After a preliminary screening, the squaraine (S6) bearing bromine as a substituent and C4 aliphatic chains showed the highest fluorescence turn-on and highest affinity for mucin than albumin. To further highlight the selectivity of S6 for mucin, the fluorescence response has been evaluated in the presence of serum and site-specific proteins different than albumin. Absorption spectroscopy was used to characterize the binding mechanism of squaraine to mucin.

INTRODUCTION

Nowadays, modern medicine firmly counts on disease detection by means of molecular biomarkers. Biomarkers are essential components in diagnosing a disease or pathogenesis process, monitoring patients during care and measuring patient pharmacological response to a therapeutic intervention (1). So far, disease diagnosis relies on the detection of biomarkers in saliva, urine, breath and seminal fluid, but the most important is based on the detection of chemicals or biological substances in the blood (2). Proteins are particularly useful molecules to use as biomarkers as they are often the effectors of diseases and the targets of pharmacological treatments. Using arrays of protein biomarkers, clinicians are able to perform accurate disease diagnoses through convenient noninvasive testing (3). Among many proteins, mucins raised great attention.

Mucins are a family of high molecular weight, rod-like and heavily glycosylated proteins. They build up the skeleton around which mucus organizes. In humans, mucins are produced and secreted by goblet cells and submucosal glands or are tethered to cell membranes (4). The main properties attributed to mucins include barrier properties, dynamicity, hydration, lubrication and bioactivity (5). In physiological conditions, epithelial cells express mucins in response to particularly harsh environments, such as the low pH in the stomach, the exposure to air in the lungs and on the eyes, and the tough environment of the intestine. Over the last decades, mucins have also been given the status of biomarkers of adverse prognosis becoming an attractive therapeutic target. Gastric carcinoma, pancreatic, colon and rectal, breast, and ovarian cancer have been investigated looking for the roles played by mucins (6–9). The overexpression of mucins in cancer cells is thought to have a similar origin as for healthy cells, namely favoring growth and proliferation in inhospitable conditions. Mucins provide a tumor-friendly environment by reducing hypoxia, acidic conditions and avoiding the immune and chemotherapy response (9). Large amounts of transmembrane mucins have been observed in mucosal epithelial cancer, or adenocarcinomas, which are associated with a poor survival rate. MUC1 is so deeply intertwined with cancer development that nowadays is recognized as an oncogene implicated in the invasion, angiogenesis and metastasis (10). In pancreatic cancer (the third most deadly type of cancer globally), mucins have been proposed as a promising biomarker for diagnosis, prognosis and therapy (11).

The last few decades have witnessed impressive advances in the development of new methods for diagnosing and monitoring diseases *via* analysis of biomarkers in body fluids that are moving in and/or out of the body (2). Several methods are currently used for mucin detection, including chemical stains or lectins, antibody-based enzyme-linked immunosorbent assays and aptamer-based electrochemical techniques (12–15). Yet, these

*Corresponding author email: nadia.barbero@unito.it (Nadia Barbero)
†This article is part of a Special Issue celebrating the 50th Anniversary of the American Society for Photobiology.

methods require specialized and expensive reagents, lengthy sample preparation procedures, challenging purification processes as well as costly equipment. In addition, multiple detection methods are recommended to have accurate identification. For these reasons, current methods can be considered time-consuming, expensive and thus unsuitable as a high-throughput routine assay. By contrast, fluorometric assays are relatively simple, straightforward, extremely sensitive and suitable for noninvasive monitoring of biological samples. To maximize fluorometric applicability in biomedical fields, a careful design and selection of selective fluorescent probes are required. Particularly, fluorescent probes with absorption and emission in the near-infrared (NIR) region (650–900 nm) garnered significant attention for biological applications since NIR signal detection does not suffer from self-absorption and autofluorescence typical of biological matrices.

In this context, squaraine dyes (SQ) are characterized by intense absorption and fluorescence emission in the visible and infrared region in organic media, which drastically decrease in aqueous media and in other nonsolvents due to the formation of aggregates that lead to fluorescence quenching. This drawback strongly limited their biological applications. Yet, this drawback can be exploited as an advantage. Indeed, without requiring the formation of covalent derivatization, some squaraines exhibit an increase in fluorescence intensity (*i.e.* turn-on) by simply mixing them with proteins. In these cases, the selective binding of SQs to specific proteins is mediated by a set of noncovalent, supramolecular interactions at specific binding sites/pockets of the protein, or simply by the variation of the surrounding environment of the SQs (16,17). The protein-induced increase in fluorescence fosters the development of squaraines as fluorescent probes of disease-related biomarkers.

Having in mind this goal, in the last few years, our group has investigated the interaction of SQ with several proteins by means of different approaches. In particular, a preliminary fluorometric assay for the detection of proteins has been optimized to identify the molecular descriptors that govern the fluorescence turn-on phenomenon. We first investigated the squaraines' fluorescence turn-on effect induced by the interaction with porcine gastric mucin (PGM) using several squaraines (*i.e.* **S1–S4**, Figure S1) bearing different functional groups (*i.e.* -COOH, −benzo-COOH) and different length of the alkyl chains (18). Increasing the concentration of PGM resulted in a gradual enhancement of fluorescence intensity. The squaraine with the higher turn-on effect (*i.e.* **S3**) was used to detect PGM in deproteinized human serum samples spiked with different amounts of mucin. Analogous turn-on tests were performed with bovine serum albumin (BSA) (16). The fluorescence turn-on between SQ with carboxylic functional group and BSA was similar, mostly higher with respect to the signal recorded in the presence of PGM. The well-known binding site of acids present in albumin may play a pivotal role in binding carboxylated squaraines such as **S1–S4**. Overall, these results indicate that the squaraines investigated so far are more specific for albumin than mucin. Next, we also investigated the binding of the same four squaraines to different proteins (*i.e.* transferrin, fibrinogen, trypsin, pepsin and a generic protease) pointing out the protein hydrophobicity as one of the key molecular properties influencing the turn-on phenomenon (19). Interestingly, we observed that squaraines with higher lipophilicity had a higher turn-on response in the presence of PGM than more hydrophilic dyes. We speculate that the squaraine aggregates may entangle with the hydrophobic domains of mucin. Once the contact is established, single dye molecules may be released from the aggregate to freely interact within the hydrophobic domains of the protein. Therefore, to increase selectivity toward mucin, more lipophilic squaraines may be necessary. Moreover, we hypothesize that removal of the hydrophilic carboxylic groups could reduce selectivity for albumin while increasing the selectivity for mucin.

In this contribution, we investigate our hypothesis by studying the turn-on response of three zwitterionic squaraines after the addition of mucin and albumin. Following a rational design, we modified the structure of the carboxylic SQ with the C2 alkyl chain (**S1**) to synthesize three new squaraines having different lipophilicity (**S5–S7** in Fig. 1). At first, we synthesize **S5** by removing the carboxylic groups on **S1** and introducing C4 aliphatic chains, obtaining a zwitterionic and more lipophilic structure. Then, the carboxylic groups were replaced with Br (**S6**) or benzo-Br (**S7**) groups, maintaining the C4 aliphatic chains, and obtaining more lipophilic structures. The interactions between these new sets of squaraines and PGM have been investigated using fluorescence spectroscopies. Spectroscopic results were compared with human serum albumin (HSA) as a reference control. After a preliminary

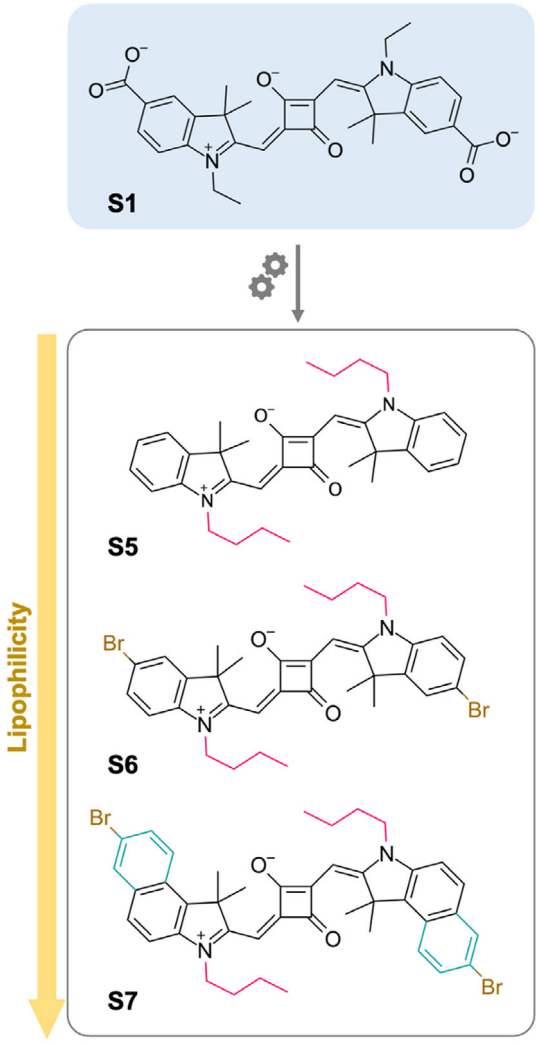


Figure 1. Structures of the squaraines tested. Molecular structures of the dicarboxylic indolenine squaraine (**S1**) (16,18,19), and the derived zwitterionic structures herein tested (**S5**–**S7**).

screening, the interaction mechanism of the squaraine having the highest selectivity for PGM was also investigated by absorbance spectroscopy and circular dichroism. In addition, the potential to probe mucin by the selected dye was evaluated by comparing the turn-on induced by other serum and site-specific proteins.

MATERIALS AND METHODS

Materials. Unless otherwise stated, all materials and reagents were purchased from commercial suppliers and used without further purification. Mucin from the porcine stomach (PGM Type III, bound sialic acid 0.5–1.5%), HSA, pepsin from porcine gastric mucosa, fibrinogen from human plasma, human transferrin, lipase from porcine pancreas, trypsin from bovine pancreas, protease from bovine pancreas, sodium chloride (NaCl) and dimethyl sulfoxide (DMSO) were all purchased from Merck. Millipore grade water was obtained from an in-house Millipore system (resistivity, 18.2 MΩ × cm at 25°C). Squaraine **S6** and **S7** were synthesized as previously described (20,21). All the chemicals for the synthesis of **S5** were purchased from Sigma Aldrich, Fluka, Merck or Riedel de Haen and were used without any further purification. All microwave reactions were performed in single-mode Biotage Initiator 2.5. TLC controls were performed on silica gel 60 F254 plates. ESI-MS spectra were recorded using an LTQ Orbitrap (Thermo Scientific) spectrometer, with an electrospray interface and ion trap as a mass analyzer. The flow injection effluent was delivered into the ion source using nitrogen as sheath and auxiliary gas. ^{1}H NMR (600 MHz) and ^{13}C NMR (151 MHz) spectra were recorded on a Bruker Avance 600 NMR in DMSO-D^{6}.

Proteins were dissolved in 0.9% NaCl at a concentration of 1 mg mL^{-1} immediately before usage. The dispersion was facilitated by mild vortexing the protein samples for 1 min. Squaraine stock solutions were prepared in DMSO at a concentration of 0.5 mg mL^{-1}. All further dilutions of proteins and of squaraines were performed in 0.9% NaCl.

Synthesis. Squaraine **S6** and **S7** were synthesized as previously described (20,21). Squaraine **S5** has been synthesized by modifying the procedure previously reported by Matsui *et al.* (22). Specifically, compared with the conventional method reported by Matsui *et al.* (22), **S5** has been synthesized by following a microwave-assisted method that allows a faster preparation of the compound, reducing the reaction time from days to minutes with more than two-fold improvement in product yields. The details of the modified procedure are reported in the following paragraphs.

Quaternization synthesis of indolenine. 1-butyl-2,3,3-trimethyl-3H-indol-1-ium iodide (**2** in Scheme 1). 2,3,3-trimethyl-3H-indole (**1** in Scheme 1) (750 mg, 4.7 mmol), iodobutane (2.14 mL, 18.8 mmol) and anhydrous acetonitrile (2 mL) were introduced in a reaction vial, sealed with a crimp cap and heated in a microwave system at 155°C for 15 min. The solvent was removed, and the solid was washed three times with diethyl ether and filtered, giving the title compound as a brownish solid (1.23 g, yield = 76%).

^{1}H NMR (600 MHz, DMSO-D^{6}) δ 7.99–7.95 (m, 1H), 7.84 (dd, J = 5.9, 2.7 Hz, 1H), 7.62 (dd, J = 5.5, 2.8 Hz, 2H), 4.45 (t, J = 7.7 Hz, 2H), 2.84 (s, 3H), 1.82 (dd, J = 9.1, 6.4 Hz, 2H), 1.53 (s, 6H), 1.43 (dd, J = 15.4, 7.5 Hz, 2H), 0.94 (t, J = 7.4 Hz, 3H).

Synthesis of symmetrical squaraine S5. (E)-4-((1-butyl-3,3-dimethyl-3H-indol-1-ium-2-yl)methylene)-2-(((E)-1-butyl-3,3-dimethylindolin-2-ylidene)methyl)-3-oxocyclobut-1-en-1-olate (**S5**). 1-butyl-2,3,3-trimethyl-3H-indol-1-ium iodide (2) (500 mg, 1.45 mmol), 4-dihydroxycyclobut-3-ene-1,2-dione (3) (83 mg, 0.73 mmol) and a mixture of toluene and *n*-butanol (1:1.5 mL) were introduced inside a sealed MW vial (2–5 mL) and heated up to 160°C for 15 min. After solvent evaporation, the crude product was crystallized by *n*-butanol, filtered off and washed with cold diethyl ether to give squaraine **S5** as golden-green crystals (350 mg, yield = 94%).

^{1}H NMR (600 MHz,) δ 7.34 (dd, J = 7.4, 0.7 Hz, 2H), 7.29 (td, J = 7.7, 1.2 Hz, 2H), 7.13 (t, J = 7.4 Hz, 2H), 6.97 (d, J = 7.9 Hz, 2H), 5.96 (s, 1H), 3.98 (s, 4H), 1.84–1.74 (m, 16H), 1.45 (dd, J = 15.4, 7.5 Hz, 4H), 0.98 (t, J = 7.4 Hz, 6H) ppm (Figure S2).

^{13}C NMR (151 MHz,) δ 170.30, 169.43, 142.55, 142.33, 127.83, 123.79, 122.37, 109.48, 86.67, 49.38, 43.63, 29.20, 27.10, 20.45, 13.95 ppm (Figure S3).

HRMS (ESI) ($C_{34}H_{40}N_2O_2$) calcd for $[M + H]^+$ 509.3163, found 509.3159.

Spectroscopic measurements. Fluorescence emission spectra in steady-state mode were acquired at room temperature using a Horiba Jobin Yvon Fluorolog 3 TCSPC fluorometer equipped with a 450-W Xenon lamp and a Hamamatsu R928 photomultiplier. Emission spectra were recorded by exciting squaraines at the wavelength corresponding to the hypsochromic shoulder observed in absorption spectra (λ_{ex} = 585 nm, λ_{ex} = 600 nm, λ_{ex} = 625 nm for **S5**, **S6** and **S7,** respectively). Initially, a constant concentration of squaraine (10 nM) was monitored in the presence of increasing concentrations of PGM or HSA. The dissociation constants (K_D) of the **S6**-PGM / HSA complex were obtained by fitting the turn-on signal (F/F_0) against the molar concentration of the protein. Data were fitted using the following nonlinear model:

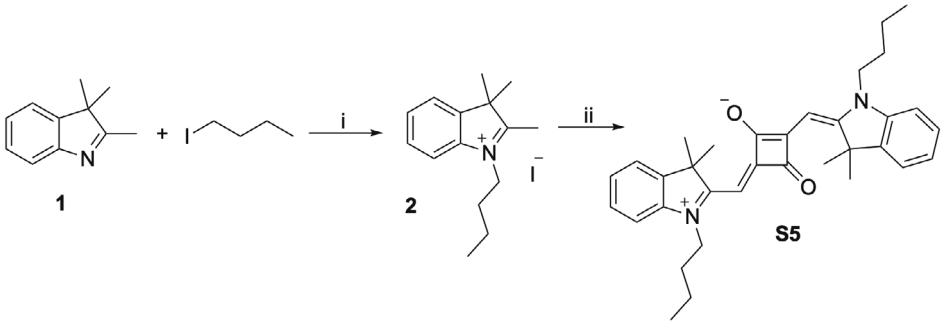
$$F/F_0 = \frac{F_{max}[P]}{K_D + [P]}$$

where F and F_0 are the fluorescence of **S6** with and without protein, respectively, F_{max} is the maximum increase in fluorescence achieved by the **S6**-protein complex at saturation, [P] is the molar concentration of the protein, while K_D is the dissociation constant in molar units.

Selectivity tests were performed by measuring the fluorescence turn-on of a 10 nM solution of **S6** in the presence of a fixed concentration (312.5 nM) of each one of the proteins tested. Spectra were acquired over the spectral range 605–800 nm, 620–800 nm and 645–800 nm for **S5**, **S6** and **S7,** respectively. Each experiment was performed in triplicate (n = 3).

Absorption spectra were measured using a UH5300 Hitachi spectrophotometer at room temperature. The absorption spectrum of a constant concentration of **S6** (2 μM) was recorded in the presence of increasing concentrations of PGM or HSA (0–300 μg mL^{-1}). Spectra were acquired at room temperature in the range 450–800 nm.

Circular dichroism measurements were performed using a Jasco J-815 CD spectrophotometer. The spectra were collected in a range of 190–260 nm in a quartz cuvette with a 0.5 mm light path using a scan speed



Scheme 1. Synthesis of squaraine **S5.** (i) acetonitrile, microwave heating at 155°C for 15–30 min. (ii) squaric acid, toluene/BuOH (1:1), microwave heating at 160°C for 30 min.

of 20 nm min^{-1} and 1 nm bandwidth. Each spectrum is the average of three scans. The concentration of protein was 100 μg mL^{-1} both in the absence and in the presence of 200 nM **S6**.

RESULTS

The brominated indolenine S6 squaraine has higher selectivity for PGM than for HSA

Symmetrical squaraine with a C4 alkyl chain (**S5**) was successfully synthesized by a multistep procedure. The synthesis of symmetrical SQ involved the condensation of the quaternary heterocyclic salts, bearing an activated methyl group, with the squaric acid. The quaternization of the 2,3,3-trimethyl-3H-indole (commercially available) to obtain compound **2** was performed under microwave irradiation, leading to an increase in the methyl group acidity which enabled the following condensation reaction (Scheme 1) to obtain **S5**. The symmetrical squaraine dye **S5** was synthesized in a one-step reaction under microwave heating following the well-established method for indolenine-based squaraines previously reported by the authors by reacting two equivalents of quaternary heterocyclic salts with squaric acid (Scheme 1) (23). The solvent of the reaction mixture was then evaporated, and the crude product was crystallized by *n*-butanol, filtered off and washed with cold diethyl ether. Squaraine **S5** was finally obtained in very high yield as golden-green crystals.

After removal of the carboxylic groups on the lateral moieties of **S1** and the addition of C4 alkyl chains, the lipophilicity of **S5** strongly raised (clogP 3.34 *vs* 5.96, calculated with MarvinSketch 20.20.0, ChemAxon). **S5** is almost nonfluorescent in 0.9% NaCl when excited at 585 nm because of the formation of insoluble aggregates in aqueous solution. Yet, with the addition of PGM, the emission intensity of **S5** increases and shows a bathochromic shift (from 636 to 646 nm). The turn-on of fluorescence induced by PGM is ~13-fold. Since albumin is the most abundant serum protein (30–50 mg mL^{-1}, approximately 60% of total serum proteins), it could induce strong background signals when trying to detect mucin at the serum level. Thus, we used HSA as a benchmark in our experiments. The fluorescence intensity of **S5** in the presence of PGM is similar in terms of turn-on and bathochromic shift (from 636 to 645 nm) to the signal resulting from the addition of HSA (Fig. 2a). These results suggest that **S5** squaraine is not specific for PGM; therefore, it cannot be used as a fluorescent probe to detect mucin at serum level. Even if there is an evident lack of selectivity for PGM, it is worth underlining that an interesting result has been achieved by the **S5** squaraine. The elimination of the carboxylic groups on the lateral moieties and the simultaneous introduction of C4 alkyl chains led to an evident reduction of the gap between the turn-

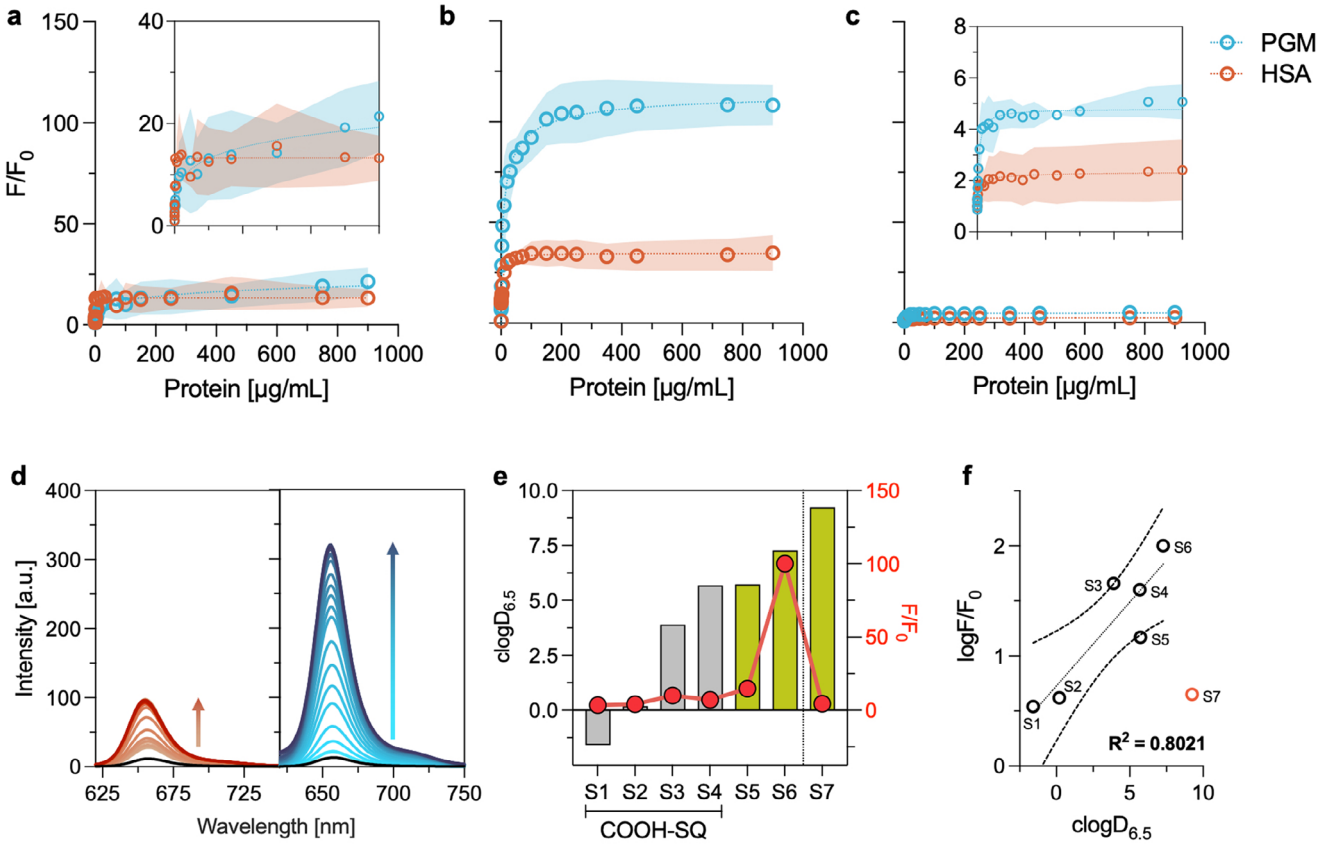


Figure 2. The brominated indolenine squaraine **S6** is more specific for PGM than for HSA. (a–c) Titration experiments of **S5**, **S6**, and **S7** respectively, with PGM and HSA. The relative increase in fluorescence is expressed as F/F_0 (turn-on). The insets in (a) and (c) represent a magnification of the respective plots. (d) Steady-state fluorescence spectra of **S6** in the presence of the same concentrations (0–900 μg mL^{-1}) of HSA (red) and PGM (blue), respectively. (e) Relation between the lipophilicity of squaraines and the fluorescence turn-on induced by PGM. Lipophilicity is expressed as the distribution coefficient at pH 6.5. A color code is used to distinguish between squaraines previously reported in the literature (**S1**–**S4**) and the squaraines herein investigated (**S5**–**S7**). (f) Correlation between the calculated distribution coefficient at pH 6.5 ($clogD_{6.5}$) of squaraines and the logarithm of the turn-on expressed as $logF/F_0$.

on response to PGM and to HSA compared with the carboxylated squaraines (**S1–S4**) that were previously investigated by our group (16,18).

It is well-known that mucin glycoproteins carry both negative charges and polar moieties as well as hydrophobic domains (24). In particular, the hydrophobic domains showed to bind with high affinity and reduce mucosal diffusion of lipophilic molecules (24–28). Following this evidence, we decided to test the effect of squaraine's lipophilicity over the interaction with PGM. Among all the squaraines tested by our group, **S5** has shown to be promising in terms of selectivity for PGM (*i.e.* similar turn-on response to PGM and HSA). Because of this, we decided to modulate the lipophilicity of **S5** with the purpose to increase selectivity toward PGM. Using **S5** as reference structure, we synthesized an analogous squaraine bearing halogenated bromo-indolenine moieties (**S6**, Fig. 1) following a previously published protocol (20). The introduction of two bromine atoms strongly raised the lipophilicity of **S6** if compared with **S5** (clogP 7.50 *vs* 5.96). To evaluate if the hypothesis of increasing the squaraines lipophilicity leads to an increase in the selectivity for PGM, we measured the fluorescence turn-on of **S6** in the presence of increasing concentrations of both PGM and HSA (Fig. 2b). As expected, the fluorescence turn-on induced by PGM was ~7 times higher than that measured with **S5** in the same conditions of concentrations. Most importantly, for the first time, we were able to record higher fluorescence signals in response to PGM than to HSA (Fig. 2b,d). The higher fluorescence signal also reflects a higher affinity for PGM with respect to HSA (1.8 *vs* 4.0 nM).

Although **S6** aggregates bind PGM with high affinity, they can still bind to HSA even though with a lower degree of affinity. This can be a relevant issue if, for example, the purpose is to detect mucin at the serum level where the concentration of albumin can exceed by several orders of magnitude the concentration of mucin. Aiming at maximizing selectivity for PGM while minimizing the response to HSA, in the next step, we sought to test whether increasing, even more, the lipophilicity of squaraine results in an increase in PGM selectivity. To obtain a structure of higher lipophilicity than **S6**, we synthesized **S7** following a protocol previously reported (21). Compared with **S6**, **S7** bears bulkier moieties (*i.e.* benzoindolenine) which raise the clogP of **S7** up to 9.48. However, the rise in lipophilicity was not followed by a simultaneous increase in selectivity for PGM. In fact, the turn-on in the presence of PGM resulted to be strongly below expectations. We recorded a ~20-fold decrease in the turn-on of fluorescence compared with **S6** (Fig. 2c). Similarly, the response to albumin was weak as well. These results could be explained by the extremely low solubility of **S7** in aqueous media. The aggregates of **S7** formed in water may be too large to ensure complete and optimal interaction either with PGM or with HSA. These results suggest that lipophilicity may play an important role in the interaction with PGM, yet up to a critical point, over which the fluorescence response does not correlate with squaraine's lipophilicity anymore. This scenario is well depicted when looking at the distribution coefficient of the tested squaraines at pH 6.5 ($logD_{6.5}$) (Fig. 2e), which is the pH of PGM suspensions when dissolved in water. In fact, due to the presence of sialic acids (29), mucin can modify the pH of not buffered solutions, shifting toward a slightly acidic environment. At this pH, a good correlation ($r^2 = 0.80$) is observed between $clogD_{6.5}$ and $logF/F_0$ (Fig. 2f).

Mucins induce the highest fluorescence turn-on of S6 squaraine

S6 squaraine has shown to be the dye with the highest fluorescent turn-on and the highest affinity for PGM. Although albumin is the most abundant protein in serum, there are plenty of other serum proteins and site-specific proteins deserving attention

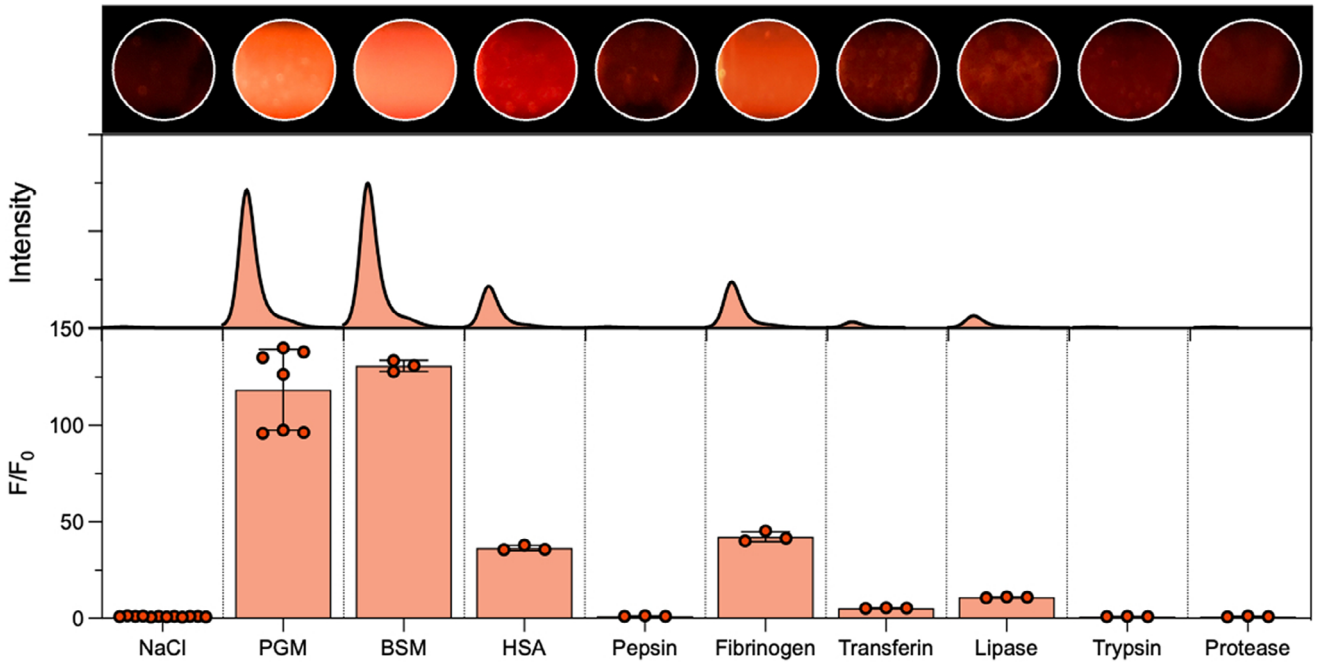


Figure 3. Gastric and oral mucins induce the highest fluorescence turn-on of **S6** squaraine. Fluorescence spectra of **S6** with different proteins and the respective maximum intensity expressed as F/F_0. Within circles, the real-time images of the quartz cuvette visible detection are reported.

when evaluating **S6** selectivity. Therefore, we evaluated the fluorescence response of **S6** to fibrinogen and transferrin as representative of serum proteins different from albumin; pepsin, lipase, trypsin and a generic protease as examples of site-specific proteins (Fig. 3). In order to compare the fluorescence response, proteins were tested at the same concentration of the previously investigated PGM. We observed that only fibrinogen, lipase and transferrin induced an enhancement of the fluorescence intensity of **S6**. In particular, an increase in fluorescence of ~42-, ~11- and ~5-fold was recorded for fibrinogen, lipase and transferrin, respectively. No increase was observed in the presence of pepsin, trypsin and protease. Yet, the fluorescence response to PGM is still the highest recorded (~120-fold increase), confirming the high selectivity of **S6** for mucin.

The mucin family includes at least 21 mucin-type glycoproteins, all of them belonging to the MUC gene. Each of these mucins has a unique structure that can influence its localization and function (30). Because of this, we investigated whether **S6** presents mucin site-specific selectivity by testing the fluorescence response to bovine submaxillary mucins (BSM). Gastric mucins (*i.e.* PGM) consist mainly of the secreted mucins MUC5AC and MUC6, produced from the superficial mucosa and the gland mucosa, respectively (31,32). Salivary mucins (*i.e.* BSM), on the contrary, consist of MUC7, MUC19, MUC1, MUC4 and MUC5B, the latter being the predominant gel-forming mucin in the oral cavity. In Fig. 3, we see that the fluorescence intensity recorded in the presence of BSM is comparable with PGM. These findings indicate that **S6** cannot discriminate between mucins having a different localization, thus suggesting that the mechanisms of interaction with PGM and BSM might be similar or the same.

Fluorescence increases following a protein-induced breakdown of S6 squaraine aggregates

Thereafter, we dove deeper into the binding between **S6** squaraine and PGM aiming at unveiling the mechanisms governing the interaction. Since SQ are sensitive to environmental changes, we next investigated the interaction between **S6** and PGM by means of UV–Vis absorption spectroscopy. The absorption of **S6** in the visible range of the electromagnetic spectrum was monitored in the presence of increasing concentrations of PGM. The spectrum of **S6** in an aqueous solution is characterized by two wide absorption bands at 600 and 650 nm: the first band being a mark of formed aggregates, while the latter represents the monomeric form. Upon the addition of PGM, we observe an increase in absorbance at 650 nm and a simultaneous bathochromic shift (600–608 nm) of the band at 600 nm (Fig. 4a–c).

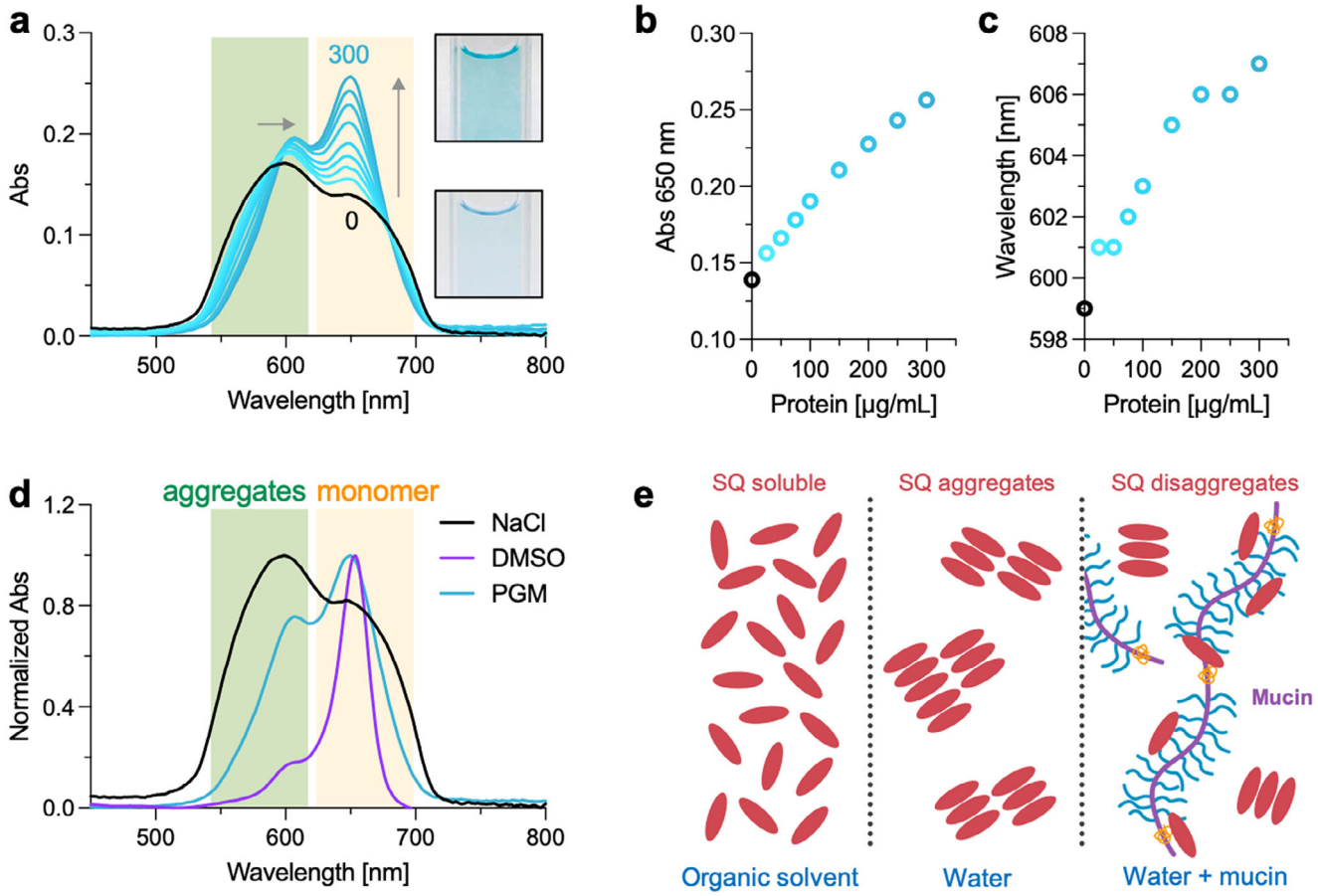


Figure 4. Mucin induces disaggregation of **S6** squaraine aggregates. (a) Absorbance spectra of **S6** (2 µM, black line) with the increasing addition of PGM. The insets display the visible detection of color saturation of squaraine samples without and with PGM. (b) Absorbance variation of the band at 650 nm while increasing the concentration of PGM (c) Bathochromic shift of the band at 600 nm while increasing the concentration of PGM. (d) Comparison between the absorption spectrum of **S6** in different environments. (e) The proposed mucin-dependent phenomenon at the origin of the recovery of the spectroscopic properties of squaraine.

Moreover, the presence of a well-defined isosbestic point at 680 nm and a second nonperfect point at ~595 nm suggests multiple cooperative equilibria, thus more than one type of binding (33,34). In addition, to clarify whether **S6** can induce a modification of the secondary structure of PGM, circular dichroism (CD) has been investigated. As reported in Figure S4, the CD spectra of PGM and PGM-**S6** complex perfectly overlap, suggesting that no variation of the secondary structure of the protein occurs.

Squaraine dyes, like most fluorescent probes, are freely soluble in organic solvents such as DMSO (35). In DMSO, **S6** appears as the typical blue color with a sharp and intense absorption band centered around 650 nm, indicative of the presence of mostly monomeric form (*i.e.* nonaggregated) (Fig. 4d). The upward absorbance at 650 nm observed with the rising concentration of PGM suggests that mucin may induce breakdown and consequently disaggregation of squaraine aggregates. The disaggregation phenomenon is the reverse process of aggregation, in which the aggregates dissociate into monomers, resulting in the enhancement and recovery of the readout signal (*e.g.* absorbance and fluorescence) (36). Indeed, previous studies revealed the capability of mucins to bring various organic and inorganic water-insoluble compounds into aqueous solution. Hendler *et al.* reported the construction of protein luminescent films by exploiting mucin's ability to keep separated organic dyes in solution (37). Another study reported the impressive and unique capability of mucin to bind and solubilize inorganic and water-insoluble nanomaterials such as C60 fullerene, fullerene-like and multiwalled carbon nanotubes (38). Our findings confirm this concept suggesting the disaggregation event as the origin of the increased spectroscopic properties of squaraines (Fig. 4e).

To identify the features of mucin that contribute to its interaction with squaraine, we compared the absorption spectra of **S6** with PGM and with carboxymethylcellulose (CMC). Similarly to mucin, CMC is a polymer that is composed of anionic and hydrophobic domains; yet, it lacks mucin-specific features such as the protein backbone and the O-linked glycans (39). Absorption spectrum of **S6** with CMC almost completely overlaps the spectrum recorded in NaCl, suggesting that **S6** aggregates do not interact with CMC (Figure S5). Instead, likewise, to what was observed with PGM, the addition of HSA slightly increases absorbance at 650 nm. However, the effect is much less pronounced with respect to PGM (Figure S5). Overall, these observations suggest that the breakdown of squaraine aggregates is not dependent on the charge but rather on the specific conformation of the protein backbone.

CONCLUSIONS

The fluorometric profiling of disease-associated biomarkers mediated by noncovalent fluorescent probes has undoubtedly gained momentum in recent years because of its unique advantages. Thus, ultrasensitive and highly selective fluorogenic probes for different biomarkers are increasingly sought-after. In this study, we make a step forward toward the development of a specific fluorescent probe for mucin. In particular, we investigated the role of lipophilicity in the interaction between three SQ with PGM. We followed a rational design to synthesize squaraines having different lipophilicity which we used to investigate the fluorescence turn-on upon interaction with PGM. All three dyes showed an increase in fluorescence intensity after the addition of PGM. Yet, only a brominated indolenine squaraine (**S6**) showed excellent fluorescence turn-on (~100-fold increase) in the presence of PGM. The same squaraine was more specific for mucin compared with other serum and site-specific proteins. By absorbance spectroscopy, we showed that mucin is able to break down the squaraine aggregates as the absorbance of the monomeric form of **S6** proportionally increases with the concentration of PGM. The same phenomenon is not manifested in the presence of carboxymethyl cellulose, a polymer comparable with mucin because of the anionic and hydrophobic domains but it is barely visible in the presence of albumin. These results suggest **S6** squaraine as a promising probe for mucin profiling. Still, many challenges remain in designing a specific fluorescent probe that responds only and uniquely to mucin. This work is expected to unravel basic principles of the mechanism of squaraine-mucin interaction and encourage researchers to develop more specific probes for mucin.

Acknowledgements—This work was supported by the University of Torino (Ricerca Locale ex-60%, Linea A, Bando 2021) and by the Fondazione CRT (II tornata 2019 RF.2019.2260). Open Access Funding provided by Universita degli Studi di Torino within the CRUI-CARE Agreement.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article:

Figure S1. Structures of the squaraines previously investigated by our group.

Figure S2. ^{1}H NMR of squaraine **S5**.

Figure S3. ^{13}C NMR of squaraine **S5**.

Figure S4. CD spectra of PGM in the absence and the presence of squaraine **S6**.

Figure S5. Absorption spectra of squaraine **S6** in NaCl in the presence of either porcine gastric mucin (PGM), human serum albumin (HSA) or carboxymethylcellulose (CMC).

REFERENCES

1. Srinitha, B. (2021) Biomarkers and disease diagnosis. *Prim. Heal. Care Open Access* **11**, 375.
2. Broza, Y. Y., X. Zhou, M. Yuan, D. Qu, Y. Zheng, R. Vishinkin, M. Khatib, W. Wu and H. Haick (2019) Disease detection with molecular biomarkers: From chemistry of body fluids to nature-inspired chemical sensors. *Chem. Rev.* **119**, 11761–11817. https://doi.org/10.1021/acs.chemrev.9b00437
3. Füzéry, A. K., J. Levin, M. M. Chan and D. W. Chan (2013) Translation of proteomic biomarkers into FDA approved cancer diagnostics: Issues and challenges. *Clin. Proteomics* **10**, 13. https://doi.org/10.1186/1559-0275-10-13
4. Ma, J., B. K. Rubin and J. A. Voynow (2018) Mucins, mucus, and goblet cells. *Chest* **154**, 169–176. https://doi.org/10.1016/j.chest.2017.11.008
5. Petrou, G. and T. Crouzier (2018) Mucins as multifunctional building blocks of biomaterials. *Biomater. Sci.* **6**, 2282–2297. https://doi.org/10.1039/c8bm00471d
6. Kufe, D. W. (2009) Mucins in cancer: Function, prognosis and therapy. *Nat. Rev. Cancer* **9**, 874–885. https://doi.org/10.1038/nrc2761
7. Hollingsworth, M. A. and B. J. Swanson (2004) Mucins in cancer: Protection and control of the cell surface. *Nat. Rev. Cancer* **4**, 45–60. https://doi.org/10.1038/nrc1251
8. Bhatia, R., S. K. Gautam, A. Cannon, C. Thompson, B. R. Hall, A. Aithal, K. Banerjee, M. Jain, J. C. Solheim, S. Kumar and S. K. Batra (2019) Cancer-associated mucins: Role in immune modulation

and metastasis. *Cancer Metastasis Rev.* **38**, 223–236. https://doi.org/10.1007/s10555-018-09775-0
9. Wi, D. H., J. H. Cha and Y. S. Jung (2021) Mucin in cancer: A stealth cloak for cancer cells. *BMB Rep.* **54**, 344–355. https://doi.org/10.5483/BMBRep.2021.54.7.064
10. Behrens, M. E., P. M. Grandgenett, J. M. Bailey, P. K. Singh, C. H. Yi, F. Yu and M. A. Hollingsworth (2010) The reactive tumor microenvironment: MUC1 signaling directly reprograms transcription of CTGF. *Oncogene* **29**, 5667–5677. https://doi.org/10.1038/onc.2010.327
11. Wang, S., L. You, M. Dai and Y. Zhao (2020) Mucins in pancreatic cancer: A well-established but promising family for diagnosis, prognosis and therapy. *J. Cell. Mol. Med.* **24**, 10279–10289. https://doi.org/10.1111/jcmm.15684
12. Croce, M. V., M. T. Isla-Larriain, S. O. Demichelis, J. R. Gori, M. R. Price and A. Segal-Eiras (2003) Tissue and serum MUCl mucin detection in breast cancer patients. *Cancer Res. Treat.* **81**, 195–207.
13. Ding, Y., J. Ling, H. Wang, J. Zou, K. Wang, X. Xiao and M. Yang (2015) Fluorescent detection of mucin 1 protein based on aptamer functionalized biocompatible carbon dots and graphene oxide. *Anal. Methods* **7**, 7792–7798. https://doi.org/10.1039/c5ay01680k
14. Wang, W., Y. Wang, H. Pan, S. Cheddah and C. Yan (2019) Aptamer-based fluorometric determination for mucin 1 using gold nanoparticles and carbon dots. *Microchim. Acta* **186**, 544. https://doi.org/10.1007/s00604-019-3516-4
15. He, Y., Y. Lin, H. Tang and D. Pang (2012) A graphene oxide-based fluorescent aptasensor for the turn-on detection of epithelial tumor marker mucin 1. *Nanoscale* **4**, 2054–2059. https://doi.org/10.1039/c2nr12061e
16. Barbero, N., C. Butnarasu, S. Visentin and C. Barolo (2019) Squaraine dyes: Interaction with bovine serum albumin to investigate supramolecular adducts with aggregation-induced emission (AIE) properties. *Chem. Asian J.* **14**, 896–903. https://doi.org/10.1002/asia.201900055
17. Xie, S., A. Y. H. Wong, S. Chen and B. Z. Tang (2019) Fluorogenic detection and characterization of proteins by aggregation-induced emission methods. *Chem. Eur. J.* **25**, 5824–5847. https://doi.org/10.1002/chem.201805297
18. Butnarasu, C., N. Barbero, C. Barolo and S. Visentin (2020) Squaraine dyes as fluorescent turn-on sensors for the detection of porcine gastric mucin: A spectroscopic and kinetic study. *J. Photochem. Photobiol. B Biol.* **205**, 111838. https://doi.org/10.1016/j.jphotobiol.2020.111838
19. Butnarasu, C., N. Barbero, C. Barolo and S. Visentin (2021) Interaction of squaraine dyes with proteins: Looking for more efficient fluorescent turn-on probes. *Dye. Pigment.* **184**, 108873. https://doi.org/10.1016/j.dyepig.2020.108873
20. Serpe, L., S. Ellena, N. Barbero, F. Foglietta, F. Prandini, M. P. Gallo, R. Levi, C. Barolo, R. Canaparo and S. Visentin (2016) Squaraines bearing halogenated moieties as anticancer photosensitizers: Synthesis, characterization and biological evaluation. *Eur. J. Med. Chem.* **113**, 187–197. https://doi.org/10.1016/j.ejmech.2016.02.035
21. Chinigò, G., A. Gonzalez-Paredes, A. Gilardino, N. Barbero, C. Barolo, P. Gasco, A. Fiorio Pla and S. Visentin (2022) Polymethine dyes-loaded solid lipid nanoparticles (SLN) as promising photosensitizers for biomedical applications. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **271**, 120909. https://doi.org/10.1016/j.saa.2022.120909
22. Matsui, M., M. Fukushima, Y. Kubota, K. Funabiki and M. Shiro (2012) Solid-state fluorescence of squarylium dyes. *Tetrahedron* **68**, 1931–1935. https://doi.org/10.1016/j.tet.2011.12.067
23. Barbero, N., C. Magistris, J. Park, D. Saccone, P. Quagliotto, R. Buscaino, C. Medana, C. Barolo and G. Viscardi (2015) Microwave-assisted synthesis of near-infrared fluorescent indole-based Squaraines. *Org. Lett.* **17**, 3306–3309. https://doi.org/10.1021/acs.orglett.5b01453
24. Käsdorf, B. T., F. Weber, G. Petrou, V. Srivastava, T. Crouzier and O. Lieleg (2017) Mucin-inspired lubrication on hydrophobic surfaces. *Biomacromolecules* **18**, 2454–2462. https://doi.org/10.1021/acs.biomac.7b00605
25. Sigurdsson, H. H., J. Kirch and C. M. Lehr (2013) Mucus as a barrier to lipophilic drugs. *Int. J. Pharm.* **453**, 56–64. https://doi.org/10.1016/j.ijpharm.2013.05.040
26. Butnarasu, C., N. Barbero, D. Pacheco, P. Petrini and S. Visentin (2019) Mucin binding to therapeutic molecules: The case of antimicrobial agents used in cystic fibrosis. *Int. J. Pharm.* **564**, 136–144. https://doi.org/10.1016/j.ijpharm.2019.04.032
27. Witten, J. and K. Ribbeck (2017) The particle in the spider's web: Transport through biological hydrogels. *Nanoscale* **9**, 8080–8095. https://doi.org/10.1039/C6NR09736G
28. Butnarasu, C., G. Caron, D. Peneda Pacheco, P. Petrini and S. Visentin (2021) Cystic fibrosis mucus model to design more efficient drug therapies. *Mol. Pharm.* **19**, 520–531. https://doi.org/10.1021/acs.molpharmaceut.1c00644
29. McShane, A., J. Bath, A. M. Jaramillo, C. Ridley, A. A. Walsh, C. M. Evans, D. J. Thornton and K. Ribbeck (2021) Mucus. *Curr. Biol.* **31**, R938–R945. https://doi.org/10.1016/j.cub.2021.06.093
30. Frenkel, E. S. and K. Ribbeck (2015) Salivary mucins in host defense and disease prevention. *J. Oral Microbiol.* **7**, 29759. https://doi.org/10.3402/jom.v7.29759
31. Skoog, E. C., Å. Sjöling, N. Navabi, J. Holgersson, S. B. Lundin and S. K. Lindén (2012) Human gastric mucins differently regulate *Helicobacter pylori* proliferation, gene expression and interactions with host cells. *PLoS One* **7**, e36378.
32. Hansson, G. C. (2020) Mucins and the microbiome. *Annu. Rev. Biochem.* **89**, 769–793. https://doi.org/10.1146/annurev-biochem-011520-105053
33. Pelosi, C., F. Saitta, C. Zerino, G. Canil, T. Biver, A. Pratesi, C. Duce, D. Fessas, C. Gabbiani and M. R. Tiné (2021) Thermodynamic evaluation of the interactions between anticancer Pt(II) complexes and model proteins. *Molecules* **26**, 2376. https://doi.org/10.3390/molecules26082376
34. Thordarson, P. (2011) Determining association constants from titration experiments in supramolecular chemistry. *Chem. Soc. Rev.* **40**, 1305–1323. https://doi.org/10.1039/C0CS00062K
35. Liu, T., X. Liu, W. Wang, Z. Luo, M. Liu, S. Zou, C. Sissa, A. Painelli, Y. Zhang, M. Vengris, M. V. Bondar, D. J. Hagan, E. W. Van Stryland, Y. Fang and K. D. Belfield (2018) Systematic molecular engineering of a series of aniline-based Squaraine dyes and their structure-related properties. *J. Phys. Chem. C* **122**, 3994–4008. https://doi.org/10.1021/acs.jpcc.7b11997
36. Wang, G., W. Xu, Y. Guo and N. Fu (2017) Near-infrared squaraine dye as a selective protein sensor based on self-assembly. *Sens. Actuators B* **245**, 932–937. https://doi.org/10.1016/j.snb.2017.01.172
37. Hendler, N., B. Belgorodsky, E. D. Mentovich, M. Gozin and S. Richter (2011) Efficient separation of dyes by mucin: Toward bioinspired white-luminescent devices. *Adv. Mater.* **23**, 4261–4264. https://doi.org/10.1002/adma.201100529
38. Belgorodsky, B., E. Drug, L. Fadeev, N. Hendler, E. Mentovich and M. Gozin (2010) Mucin complexes of nanomaterials: First biochemical encounter. *Small* **6**, 262–269. https://doi.org/10.1002/smll.200900637
39. Samad, T., J. Witten, A. J. Grodzinsky and K. Ribbeck (2022) Spatial configuration of charge and hydrophobicity tune particle transport through mucus. *Biophys. J.* **121**, 277–287. https://doi.org/10.1016/j.bpj.2021.12.018