ChemComm

Cite this: *Chem. Commun.*, 2012, **48**, 6681–6683

www.rsc.org/chemcomm

COMMUNICATION

# Development of a fluorescent chalcone library and its application in the discovery of a mouse embryonic stem cell probe†

**Sung-Chan Lee,‡[a] Nam-Young Kang,‡[a] Sung-Jin Park,[a] Seong-Wook Yun,[a] Yogeswari Chandran[a] and Young-Tae Chang*[ab]**

*Received 6th March 2012, Accepted 11th May 2012*
**DOI: 10.1039/c2cc31662e**

**We report the first fluorescent diamino-chalcone library and its application in the discovery of a mouse embryonic stem cell (mESC) probe. CDg4, a novel green fluorescent mESC probe was discovered through a high-content image based screening of 160 members of the chalcone library. Interestingly, the molecular binding target of CDg4 was identified as the glycogen of the stem cell colony surface, rather than a conventional protein target from an intracellular source.**



The diversity directed probe development approach by fluorescent library synthesis and high throughput screening has been demonstrated as a powerful tool for novel sensor/probe discovery.[1] Considering the limited number of currently available fluorescent scaffolds, elucidation of novel chemical structures with desirable fluorescence properties is one of the most important components in expanding the diversity coverage of the tool boxes in novel sensor/probe discovery. Herein, we report the first diamino-chalcone based fluorescent dye library and the discovery of a novel green fluorescent probe for mouse embryonic stem cells (mESC).

Chalcones are considered to be precursors of flavonoids and isoflavonoids in nature, and have been known to possess various biological activities.[2] Besides their medicinal applications, the photophysical and photochemical properties of chalcones have been neglected for a long time, whilst some chalcone derivatives with proper electron push–pull pairs are known to be fluorescent (Fig. 1).[3]

By analyzing the structures of known fluorescent chalcone compounds, we designed a novel fluorescent ChaLcone Amide library (CLA) as shown in Fig. 1. CLA compounds are designed to have one amide and one amine on each side of the scaffold as electron donors, to enhance the fluorescence emission intensity compared to other known mono oxygen or mono amino chalcone fluorescent dyes.[3]

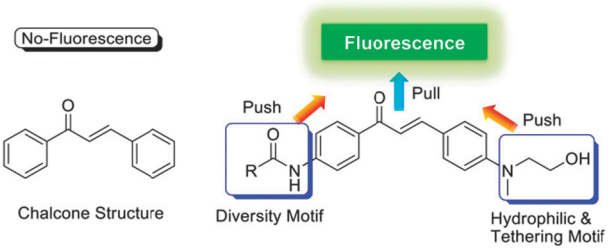
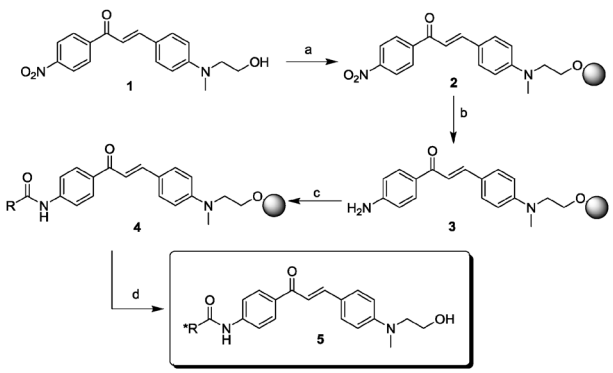



**Fig. 1** Structural design of Chalcone Library (**CLA**).

The synthesis of the CLA library is outlined in Scheme 1. Nitro chalcone **1** was prepared by microwave assisted condensation of 4-nitroacetophenone and *N*-alkyl-4-aminobenzaldehyde using pyrrolidine as a base. Intermediate **1** was loaded to 2-chlorotrityl chloride resin and the nitro group was converted to an amino group *via* $SnCl_2$ reduction. The reduced amino group was then coupled with 160 commercially available acid chlorides to generate structural diversity within the chalcone scaffold. Mild acidic cleavage with 2% trifluoroacetic acid in dichloromethane was followed by simple filtration through a silica flock to afford 160 library compounds. Without further purification, LC–MS




**Scheme 1** Synthetic scheme: (a) 2-ClTr chloride resin, pyridine, $CH_2Cl_2$, 5 h, RT; (b) $SnCl_2$, 6% $H_2O$ in DMF, RT, overnight; (c) acid chlorides, pyridine, $CH_2Cl_2$, 30 min.; (d) 2% TFA in $CH_2Cl_2$; DMF = dimethylformamide, TFA = trifluoroacetic acid, RT = room temperature, 2-ClTr = 2-chlorotrityl, *R is from 160 commercial acid chlorides (Table S1†).

[a] *Laboratory of Bioimaging Probe Development, Singapore Bioimaging Consortium, Agency for Science, Technology and Research (A*STAR), Biopolis, Singapore 138667*
[b] *Department of Chemistry, National University of Singapore, MedChem Program of Life Sciences Institute, National University of Singapore, 3 Science Drive 3, Singapore 117543. E-mail: chmcyt@nus.edu.sg*

† Electronic supplementary information (ESI) available: Experimental procedures, spectroscopic data, cell image data. See DOI: 10.1039/c2cc31662e

‡ These authors contributed equally.

analysis of the library showed high purities (average purity was 85% at 430 nm) and their spectroscopic properties were measured individually and fully documented in Table S2†. Most of the library compounds showed similar optical properties with a maximum absorbance at 430 nm, and emission at 560 nm. Quantum yields were distributed between 0.1 and 0.2 and the *c*LogP (average 4.5) were in a reasonable range for biological applications according to the Lipinski rule (Fig. S1†).[4] In rare cases, lower quantum yield ($<0.1$) was observed when strong electron withdrawing groups were introduced on the amide group. This may be attributed to the fact that the strong electron withdrawing groups attenuate the electron donor's pushing effect to the chalcone scaffold. The structure and quantum yield relationships are summarized in Table S3 and Fig. S2†.

For novel biological applications, the synthesized CLA library was screened for mESC selective probes. Using whole cell image based screening, our group has previously reported one yellow (**CDy1**) and one blue (**CDb8**) mESC probe.[5] We envisioned the green color of CLA would allow the chance of complementary multi-color imaging probe development in the green range. All the CLA compounds were screened in mESC and mouse embryonic fibroblast (MEF) as a negative control in 384 well plates. The cell images were recorded by the automated imaging microscope system, ImageXpress Micro™, and the fluorescent images were analyzed by cellular fluorescence intensity using MetaXpress® image processing software. From this initial screening, we selected 3 compounds that displayed selective staining for mESC. Subsequently, we validated the 3 candidate compounds by flow cytometry to identify **CDg4** (Compound of Designation green 4), as the only compound which works both for imaging and flow cytometry (Fig. 2).

For further characterization of **CDg4** in biological applications, we induced differentiation of **CDg4**-stained mESC by removing LIF from the culture media and allowing them to form embryoid

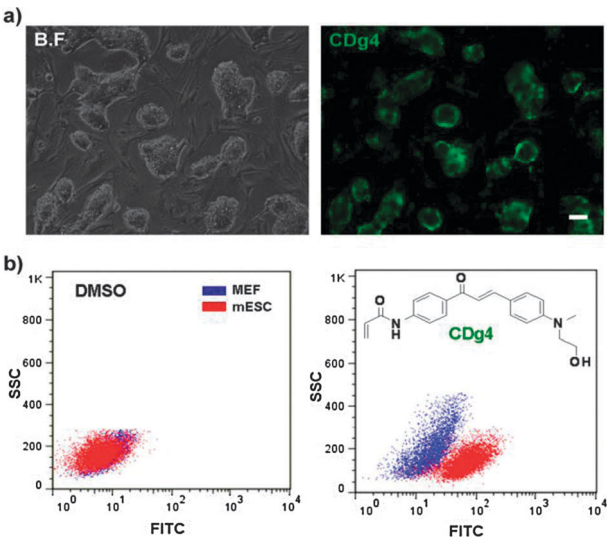



**Fig. 2** (a) The mESCs were co-cultured on MEF. The bright field (left) and fluorescent image (right) of **CDg4**. Scale bar: 200 μm. Images of stained cell colonies were taken by ×4 objective lenses. (b) Flow cytometry dot plot images of mESC and MEF stained with DMSO as a control and **CDg4** ($\lambda_{ab}$: 430 nm, $\lambda_{em}$: 560 nm, $\varepsilon$: 23 600, QY: 0.2 in DMSO).

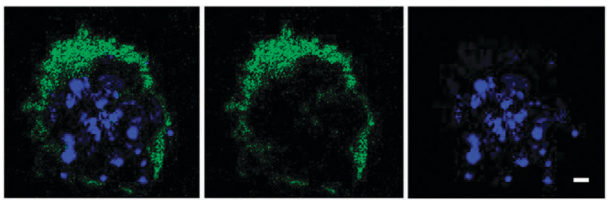


**Fig. 3** Confocal images for a mESC colony stained by **CDg4**. Merged fluorescent images of **CDg4** and Hoechst 33 342 (left), **CDg4** fluorescent image (middle: 2 μM of **CDg4** is incubated for 1 h) and nucleus staining image with Hoechst 33 342 (right). Scale bar: 10 μm. Images of stained cell colonies were taken by ×40 objective lenses.

bodies and spontaneously differentiate. The 3 types of germ layer cells, which were confirmed by immunostaining of lineage specific markers, were not stained by **CDg4** (Fig. S3†). Then we split the mESC and stained them with **CDg4** at each passage to investigate the effect of **CDg4** on mESC proliferation. The numbers of mESC colonies counted from the **CDg4**-stained groups during 3 serial passages were 56 ± 6, 56 ± 6 and 57 ± 2 per 4.05 $mm^2$, which were not statistically different from 50 ± 5, 53 ± 8 and 51 ± 7 obtained from a non-treated group (Fig. S4†). These results suggest that **CDg4** does not affect the normal growth, colony formation and differentiation of mESC. For the subcellular localization analysis of **CDg4** by multi-colour imaging, we stained mESC using **CDy1**, **CDb8**, **CDg4** and Hoechst 33 342. The multi-colour images (Fig. S5, S6†) and confocal microscopy (Fig. 3) showed that **CDg4** stains only the outside of a mESC colony.

To validate the generality of the mESC staining property of **CDg4**, we attempted to stain the E14 mESC cell line and a mouse induced pluripotent stem cell (iPSC) which was generated from the MEF of Oct4-GFP transgenic mice by ectopic expression of Oct-4, Sox2, Klf4 and c-Myc. E14 was stained by **CDg4** with a similar pattern as shown in our mESC cell line (Fig. S7†). For GFP expressing iPSC colony staining, we took the images before and after **CDg4** staining. The image taken after **CDg4** staining obviously showed larger area, which implies that the outsides of the colonies were stained by **CDg4** (Fig. S8†).

The stem cell colony surface is composed of numerous polysaccharides, lipids, lipoproteins, and glycolipids which play roles both in cell–cell interaction and embryo development.[6] To investigate the possible binding target of **CDg4**, we performed a systematic fluorescence response assay for **CDg4** against 82 biological analytes *in vitro* (Fig. S9†).[7] To our surprise, **CDg4** selectively responded to only glycogen among 82 analytes, increasing its fluorescent intensity.

The selective response of **CDg4** on glycogen brought us to a further focused screening using various structurally related carbohydrate polymers. Interestingly, **CDg4** showed almost no response to dextran, amylose, dermatan sulphate, hyaluronic acid and heparin, which are structurally similar carbohydrate polymers to glycogen (Fig. 4a), but uniquely responded to glycogen in a dose dependent manner reaching up to a 66 fold fluorescence increase. This was a striking observation, especially for dextran and amylose, which have exactly the same building blocks, *i.e.* glucose, with glycogen. Glycogen is a highly branched glucose polymer with α(1 → 4) linkage in the main chain and α(1 → 6) branch points. Amylose is a linear polymer

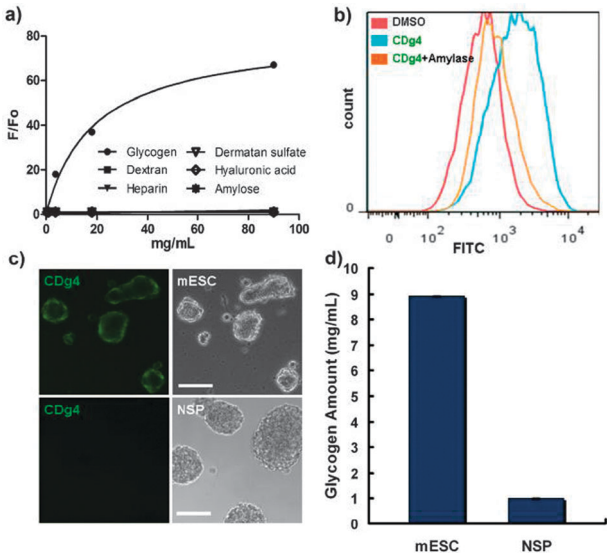



**Fig. 4** (a) Fluorescence increase of **CDg4** spectra (10 μM) upon interaction with serial concentrations of glycogen (ex: 430 nm). Fluorescent intensity was measured at 560 nm. The same concentrations of dextran, dermatan sulphate, hyaluronic acid, amylose and heparin were used. (b) Histogram of flow cytometry analysis. 0.5% amylase was treated to mESC after 2 μM of **CDg4** staining. The **CDg4** signal was measured by flow cytometry. (c) **CDg4** fluorescent images (left) and bright field (right) images of mESC and mouse neurosphere (NSP). Scale bar: 100 μm. (d) Glycogen quantification assay for mESC and neurosphere (NSP).

with only α(1 → 4) linkage, and dextran has α(1 → 6) main chains with short α(1 → 3) branches. We speculate the selective response of **CDg4** to glycogen is due to the unique secondary structure derived from the α(1 → 6) branched glucose structure. Also, glycogen is a globular shaped polymer with longer branch units and a larger exposed area, compared to dextran and amylose, which are in relatively more packed forms.[8] Glycogen is known for energy storage and the accumulation is dramatically increased during early embryonic development.[9] High glycogen accumulation in the inner cell mass (ICM) and also inter cellular localization has been demonstrated as depicted in Fig. S10†.[9c] To confirm if glycogen on the mESC colony surface is indeed responsible for the selective staining of **CDg4** to mESC in comparison to other cell types, we hydrolyzed the surface glycogen of mESC using amylase.

The amylase treatment dramatically reduced the fluorescence intensity of mESC stained by **CDg4** (Fig. 4b) indicating that the molecular binding target of **CDg4** is the glycogen located on the surface of the mESC colony. Next, we examined the **CDg4** staining selectivity in neurosphere (NSP) to rule out the possibility that **CDg4** simply stains clustering cells. It was clearly demonstrated that mouse NSP was not stained by **CDg4**, despite their similar physical morphology (Fig. 4c). To confirm the connection of this selectivity to glycogen content on the colony surface, we quantified[10] the glycogen in each type of stem cell and found that mESC has 9 times higher glycogen content than NSP (Fig. 4d). Altogether, these data suggest that glycogen is the binding target of **CDg4** in mESC. This is an intriguing example of a fluorescent small molecule targeting a carbohydrate, considering the fact that most of the conventional small molecule targets in chemical genetics are proteins.[11]

The glycogen specific probe **CDg4** may enable the detection of mESC and the quantification of embryonic surface glycogen levels.

This study was supported by intramural funding and grant 10/1/21/19/656 from the A*STAR (Agency for Science, Technology and Research, Singapore) Biomedical Research Council.

## Notes and references


1 (*a*) M. Vendrell, J. S. Lee and Y. T. Chang, *Curr. Opin. Chem. Biol.*, 2010, **14**, 383; (*b*) J. S. Lee, N. Y. Kang, Y. K. Kim, A. Samanta, S. H. Feng, H. K. Kim, M. Vendrell, J. H. Park and Y. T. Chang, *J. Am. Chem. Soc.*, 2009, **131**, 10077; (*c*) N. Y. Kang, H. H. Ha, S. W. Yun, Y. H. Yu and Y. T. Chang, *Chem. Soc. Rev.*, 2011, **40**, 3613; (*d*) R. K. Das, A. Samanta, H. H. Ha and Y. T. Chang, *RSC Adv.*, 2011, **1**, 573.

2 (*a*) S. S. Mokle, M. A. Sayyed, Kothawar and Chopde, *Int. J. Chem. Sci.*, 2004, **2**, 96; (*b*) H. K. Hsieh, L. T. Tsao, J. P. Wang and C. N. Lin, *J. Pharm. Pharmacol.*, 2000, **52**, 163; (*c*) G. S. Viana, M. A. Bandeira and F. Matos, *Phytomedicine*, 2003, **10**, 189.

3 (*a*) M. Gaber, T. A. Fayed, S. A. El-Daly and Y. S. El-Sayed, *Photochem. Photobiol. Sci.*, 2008, **7**, 257; (*b*) K. Rurack, J. L. Bricks, G. Reck, R. Radeglia and U. Resch-Genger, *J. Phys. Chem. A*, 2000, **104**, 3087; (*c*) C. G. Niu, A. L. Guan, G. M. Zeng, Y. G. Liu and Z. W. Li, *Anal. Chim. Acta*, 2006, **577**, 264; (*d*) V. Marchi-Artzner, B. Lorz, C. Gosse, L. Jullien, R. Merkel, H. Kessler and E. Sackmann, *Langmuir*, 2003, **19**, 835; (*e*) M. Ono, R. Ikeoka, H. Watanabe, H. Kimura, T. Fuchigami, M. Haratake, H. Saji and M. Nakayama, *ACS Chem. Neurosci.*, 2010, **1**, 598; (*f*) P. F. Wang and S. K. Wu, *J. Photochem. Photobiol., A*, 1995, **86**, 109; (*g*) D. G. Powers, D. S. Casebier, D. Fokas, W. J. Ryan, J. R. Troth and D. L. Coffen, *Tetrahedron*, 1998, **54**, 4085; (*h*) A. Ullah, F. L. Ansari, I. Haq, S. Nazir and B. Mirzab, *Chem. Biodiversity*, 2007, **4**, 203; (*i*) A. R. Jagtap, V. S. Satam, R. N. Rajule and V. R. Kanetkar, *Dyes Pigm.*, 2011, **91**, 20.

4 C. A. Lipinski, F. Lombardo, B. W. Dominy and P. J. Feeney, *Adv. Drug Delivery Rev.*, 2001, **46**, 3.

5 (*a*) C. N. Im, N. Y. Kang, H. H. Ha, X. Z. Bi, J. J. Lee, S. J. Park, S. Y. Lee, M. Vendrell, Y. K. Kim, J. S. Lee, J. Li, Y. H. Ahn, B. Feng, H. H. Ng, S. W. Yun and Y. T. Chang, *Angew. Chem., Int. Ed.*, 2010, **49**, 7497; (*b*) K. K. Ghosh, H. H. Ha, N. Y. Kang, Y. Chandran and Y. T. Chang, *Chem. Commun.*, 2011, **47**, 7488.

6 (*a*) A. V. Nairn, A. Kinoshita-Toyoda, H. Toyoda, J. Xie, K. Harris, S. Dalton, M. Kulik, J. M. Pierce, T. Toida, K. W. Moremen and R. J. Linhardt, *J. Proteome Res.*, 2007, **6**, 4374; (*b*) R. J. Linhardt and T. Toida, *Acc. Chem. Res.*, 2004, **37**, 431; (*c*) B. Thisse and C. Thisse, *Dev. Biol.*, 2005, **287**, 390.

7 J. S. Lee, H. K. Kim, S. H. Feng, M. Vendrell and Y. T. Chang, *Chem. Commun.*, 2011, **47**, 2339.

8 (*a*) B. J. Gibbons, P. J. Roach and T. D. Hurley, *J. Mol. Biol.*, 2002, **319**, 463; (*b*) J. Lomako, W. M. Lomako and W. J. Whelan, *Biochim. Biophys. Acta, Gen. Subj.*, 2004, **1673**, 45; (*c*) A. M. Myers, M. K. Morell, M. G. James and S. G. Ball, *Plant Physiol.*, 2000, **122**, 989.

9 (*a*) C. B. Ozias and S. Estern, *Biol. Reprod.*, 1973, **8**, 467; (*b*) W. Dormeyer, D. V. Hoof, S. R. Braam, A. J. R. Heck, C. L. Mummery and J. Krijgsveld, *J. Proteome Res.*, 2008, **7**, 2936; (*c*) S. Kolle, M. Stojkovic, K. Prelle, M. Waters, E. Wolf and F. Sinowatz, *Biol. Reprod.*, 2001, **64**, 1826.

10 F. Huijing, *Clin. Chim. Acta*, 1970, **30**, 567.

11 D. P. Walsh and Y. T. Chang, *Chem. Rev.*, 2006, **106**, 2476.