5-8-2015

Applications of Organic Probes for Imaging and Analysis of Different Cancer Cell Models

Hong Guan
University of South Carolina - Columbia

APPLICATIONS OF ORGANIC PROBES FOR IMAGING AND ANALYSIS OF DIFFERENT CANCER CELL MODELS

by

Hong Guan

Bachelor of Science
Tsinghua University, 2005

Master of Science
Medical College of Wisconsin, 2009

Submitted in Partial Fulfillment of the Requirements

For the Degree of Doctor of Philosophy in

Chemistry

College of Arts and Sciences

University of South Carolina

2015

Accepted by:

Qian Wang, Major Professor

John H. Dawson, Committee Member

Stephen L. Morgan, Committee Member

Guiren Wang, Committee Member

Lacy Ford, Vice Provost and Dean of Graduate Studies

Dedication

This dissertation is dedicated to my loving, optimistic and supportive husband, my cute smart little boy and girl, and to my always supportive and faithful parents and brother. Your love and positive energy have provided the wings to fly higher.

Acknowledgements

Great appreciation and sincere thanks for the suggestions and support are present here to the following persons who have made this dissertation possible.

Dr. Qian Wang, my advisor and my life guider, his suggestions and foresights have been beneficial not only to my research, but also to my life and my future career. I am deeply indebted to his stimulating motivation and valuable ideas which help to make this work possible. His patience and kindness have been with me for each moment in the life. His caring and support are also extended to my family. He is a great supervisor for us in every field.

Dr. Xinrui Duan, I would like to thank him for his innovation and patience with me. He has taught me how to think critically and scientifically. His curiosity and enthusiasm in science lead me to enjoy every day working.

I would like to thank my fellow lab members for their valuable comments and suggestions, including Dr. Honglin Li, Dr. Jun Hu and Dr. Yuping Shan for sharing their expertise in organic chemistry and analysis chemistry, as well as their constructive discussion and valuable comments on my research. Thanks to my lab member Xia Zhao who has been with me and engaged me with scientific talking.

Dr. Chang-UK Lim and Dr. Udai Singh, their helps in performing and analysis of the flow cytometry assay in Chapter 1 and Chapter 4 are greatly

appreciated. Their instructions and suggestions on cell analysis and sorting provided me invaluable experience.

I greatly appreciate helps and suggestions from fellows of Dr. Hexin Chen's lab, including graduate student Yogin Patel, and Dr. Shou Liu. Their suggestions and help in cell culturing and signaling pathway analysis in Chapter 1, 2 and 4 have expanded my knowledge.

I also owe thanks to Dr. Franklin G. Berger, Dr. Daping Fan for their helps on searching for proper cell lines. I would like to extend my gratitude to Dr. Andrew Greytak and Yi Shen for their collaboration in synthesis DBCO-Quantum dots in Chapter 3.

Great gratitude and sincere thanks to Dr. John H. Dawson, Dr. Stephen L. Morgan, Dr. Guiren Wang, my committee members, for reviewing and evaluating this dissertation and their invaluable suggestions and kind help throughout my graduate study.

I would like to thank my family and friends for their endless love and supports. My classmates and friends from University of South Carolina, especially Jie Gao's family and Chao Li, they always stand by me and are willing to help. I would like to extend my appreciation to my friends Yuwei Kan and Xiaofang Zhang, for being my best friends and sharing their life with me. My parents in law and my sister in law have been with me for each happy moment. Their encouragement and good wishes are meaningful to me.

I owe great appreciation and love to my special one, my husband Xiaoguang Ma, for his sweet understanding, supports and love. My little ones, Rachel and Tyler, I am so

grateful to have you two in my life. You cute faces are source of my happiness and strength.

Special thanks to my brother for sending his love over the phone. His love has securely protected me. Lastly, I would like to express my enormous thanks to my dear parents for their dedication, their hard working and their endless love. They are my source of strength, wisdom and inspiration. I am forever grateful for being their daughter.

Abstract

Novel techniques for better studying cancer cell behavior and progression are extremely important. Our lab is interested in synthesizing specifically designed organic reagents and applying them in the studies of cancers. In this dissertation, we used organic probes to investigate cancer cells in the following topics: determination of amino acid concentration in different types of cancers, tumor toxicity of organic probes towards HER2 positive breast cancer cells, glycoproteins identification in breast cancer cells and development of enzyme activity assay for colon cancer cells.

The first chapter was focused on the determination of cysteine/homocysteine concentrations using turn-on sensors in cancer cells. Traditional methods such as HPLC are time consuming and cumbersome. DCP family probes were synthesized by our group previously, of which three probes reacted with cysteine/homocysteine specifically. To test the reactivity, the sensors were reacted with different substrates, including amino acid, nucleotides and proteins. All three sensors showed great selectivity towards cysteine and homocysteine. We confirmed the correlation between fluorescence intensity and cysteine concentration. Then liver cancer cell model which was previously reported with different intracellular cysteine concentration was used to test the sensitivity of sensor **2** inside live cells. The staining results from the liver cancer cells pretreated with 400 μM of cysteine showed higher intensity of fluorescence than the cells pretreated with 0 μM.

To study cysteine/homocysteine concentration in breast cancer cells, we used sensor **2** to stain genetic modified breast cell line MCF10A *Vector* and MCF10A *HER2*, which could mimic normal breast cells and breast cancer cells. MCF10A *HER2* had higher cysteine/homocysteine level than MCF10A *Vector*.

Besides being good fluorescent probes, DCP family probes may also work as anti-tumor drugs. To study the toxicity of DCP family probes, we have performed cell viability assay to screen possible probes targeting the breast cancer marker HER2 in Chapter 2. Several probes were selected and the HER2 downstream signaling pathways were further studied. It was found that sensor **3** may be further modified for HER2 specific drugs to treat breast cancer.

In Chapter 3, we used bioorthoganol probes and correlating reporters to develop protocols for cell imaging and enrichment of glycoproteins. Using two different fluorescent reporters, we addressed the total glycosylation level as well as cell surface specific glycosylation in three different breast cell lines. We also confirmed that breast cancer cells had elevated the level of glycosylation compared to the normal breast cells in this chapter. To enrich glycoproteins, our lab modified a silicone bead with disulfide bond and triple bond at the end of surface chain. The triple bond could react with azide and be used to fish labeled proteins while the disulfide bond could be cleaved to release the proteins after enrichment. Using this probe, we tried different methods for glycoprotein purification. Unfortunately, we were not able to develop a protocol to enrich glycoproteins with these beads.

In the last Chapter, we moved to the enzyme activity detection in colon cancer cell line. A fluorescent sensor named RSAAA was synthesized specifically targeting on mitochondria and might be worked potentially as substrate for aldehyde dehydrogenase 2 (ALDH2). We tested the efficiency of this sensor working as substrate for ALDH2. Based on the results, we developed a protocol to stain cells and isolated populations with different ALDH2 activity. To study the role of ALDH2 in cancer stemness, ALDH2 high and ALDH2 low cell populations were isolated using the protocol set up in this chapter. The tumorigenicity of two populations was studied to identify the enrichment of cancer stem cells. In HCT15, ALDH2 high cells had more cancer stem cell population and were more malignant compared to the ALDH2 low population.

Specific designed probes could be powerful tools in the study of biological progress. We have devoted to the application of multiple probes to better understand and study cancer cells. Good communications between chemistry and biology help to support the translation from basic organic chemistry to the applications in biological fields. The emerging field of studies would be greatly appreciated in the development of new methods for biological processes.

Table of Contents

List of Tables

List of Scheme

List of Figures

List of Abbreviations

ABC ..... ATP-binding cassette

ALDH ..... Aldehyde dehydrogenase

BAA- ..... BODIPY-aminoacetate

BAAA ..... BODIPY-aminoacetaldehyde

Bcl-2 ..... B-cell lymphoma 2

CSC ..... Cancer stem cell

Cys ..... Cysteine

DEAB ..... Diethylaminobenzaldehyde

ELISA ..... Enzyme-linked immunosorbent assay

ErbB ..... Epidermal growth factor receptor

FBS ..... Fetal bovine serum

FGFR ..... Fibroblast growth factor receptor

GAPDH ..... Glyceraldehyde 3-phosphate dehydrogenase

GCL ..... Glutamate-cysteine ligase

GSH ..... Glutathione

Hcy ..... Homocysteine

HER2 ..... Human epidermal growth factor receptor 2

HPLC ..... High performance liquid chromatography

Mcl-1 ..... Myeloid cell leukemia sequence-1

PBS ..... Phosphate buffered saline

QRT-PCR ..... Quantitative reverse transcription polymerase chain reaction

CHAPTER 1 Highly Selective Fluorescence Turn-on Sensors for Cysteine/Homocysteine for Live Cell Imaging

1.1 Introduction

1.1.1 The concentrations of thiol containing molecules are tightly controlled in cells

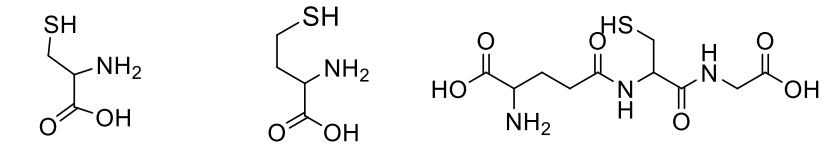
There are many different thiol containing molecules in cells, including cysteine, homocysteine and glutathione (Fig. 1.1.a), which play essential roles in human physiology [1].

Cysteine (Cys) is a sulfur-containing amino acid that can be endogenously synthesized in human cells. The cysteine metabolism is showed in Fig. 1.2. Besides supporting protein synthesis, it is also the precursor for many biomolecules like glutathione, hypotaurine, taurine and inorganic sulfates [2]. The concentration of cysteine is tightly controlled in cells. While low cysteine levels might be associated with cardiovascular disease [3] and hematopoiesis decrease [4], high concentration of cysteine was toxic to cells. For example, high levels of intracellular cysteine could induce DNA damage by driving the Fenton reaction [5]. In addition, high plasma concentration of cysteine was associated with development of rheumatoid arthritis [6], Parkinson's disease [7], Alzheimer disease [8].

Homocysteine (Hcy) is another sulfur-containing amino acid. Compared to cysteine (193 μM), homocysteine (9.3 μM) is almost negligible in cells [9]. Similar as Cys, Hcy is a risk factor for many diseases, including cardiovascular [10] and

.

a)



Cysteine (Cys) Homocysteine (HCy) Glutathione (GSH)

b)

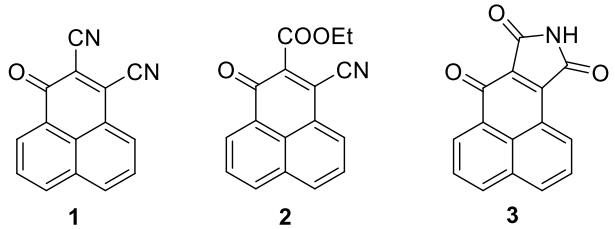


Figure 1.1. Structures of thiol containing molecules and fluorescence sensors. a) Structures of cysteine, homocysteine and glutathione. b) Structures of three fluorescence turn-on sensors used in this research.

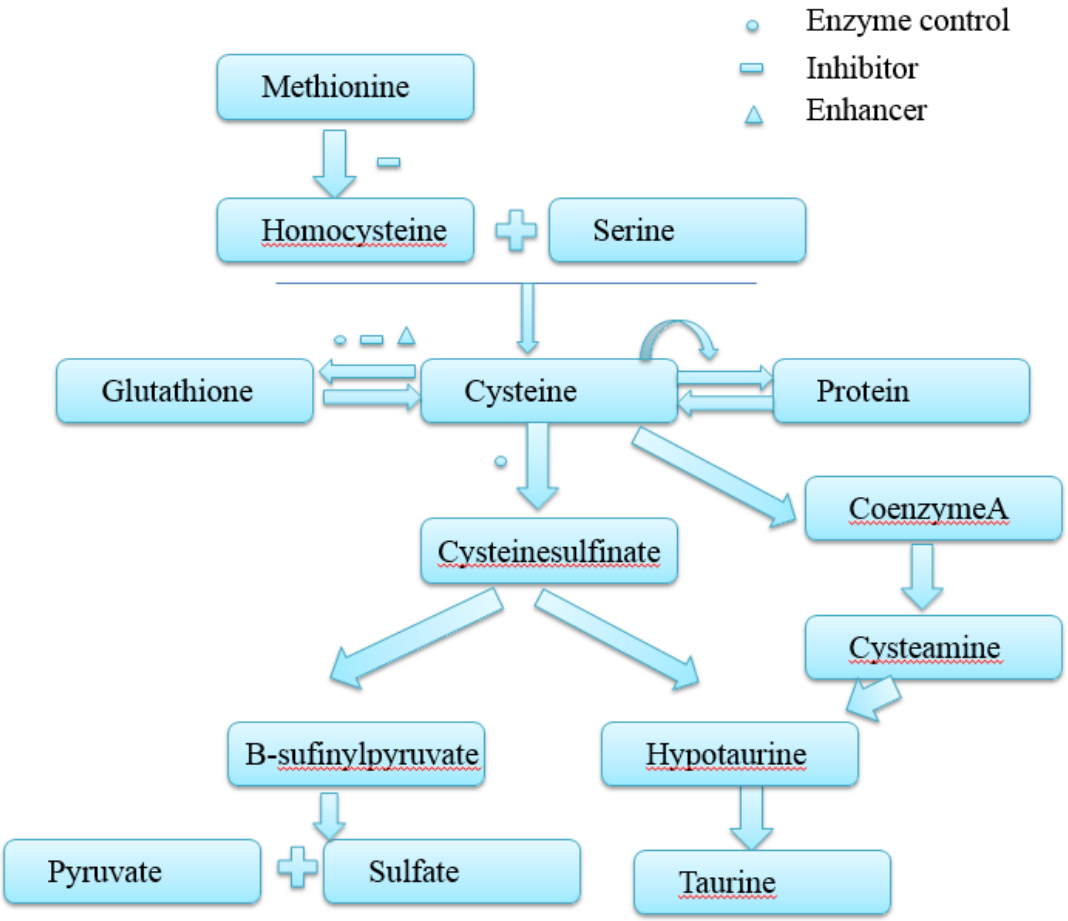


Figure 1.2. Schematic illustration of the cysteine metabolism pathways [2, 11]. Cysteine can be synthesized from homocysteine and serine and it is the precursor for glutathione. Most of cysteine contributed for the protein synthesis. At each step, circle, rectangle and tricycle are used to stand for enzymatic control, inhibitor and enhancer, respectively.

Alzheimer's disease [12]. Furthermore, total homocysteine concentration in plasma is related to birth defects and cognitive impairment in old people [13].

Glutathione (GSH) is the most abundant intracellular nonprotein-thiol (1.0-15.0 mM [14, 15]). It plays a pivotal role in maintaining a reducing environment in cells [16], acting as the redox regulator [17] and regulating the expression of different genes [18, 19]. A redox homeostasis exists between sulfhydryl and disulfide forms for GSH [19]. Glutathione has also been reported to be involved in many diseases, including cancer [20], Alzheimer's [21], and cardiovascular disease [22, 23].

Cysteine, homocysteine and glutathione metabolisms are closely correlated as shown in Fig. 1.2. For example, homocysteine is the precursor of cysteine [24], while cysteine is the precursor for the synthesis of glutathione and the rate determine reagent [25]. The enzyme, glutamate-cysteine ligase (GCL), catalyzes the synthesis of GSH from cysteine and glutamate [26]. It has been known that the Km of GCL from rat liver is 0.1 mmol/L, which makes glutathione synthesis extremely sensitive to the intracellular cysteine concentration. This process is reversible when cysteine concentration is low [2]. The relations between these three thiol-containing molecules imply important roles of cysteine in disease development, progression and treatment [10, 12, 27-29].

1.1.2 Cysteine concentration changes during diseases development

Changes of intracellular and extracellular cysteine concentrations could both work as signals for cellular responses such as gene expression, protein synthesis and oxidative stress relief [25, 30]. Recent study had demonstrated that cysteine availability inside cells could control the thiolation of tRNA to regulate protein translation and metabolism

homeostasis [31]. Several studies showed cysteine availability from neighbor cells was essential for cell survival and scavenge of oxidative stress [30, 32].

Cysteine concentration can be modulated by the transportation [33, 34] and metabolism [2] of cysteine or cystine. Cystine is the oxidized form of cysteine. Once cystine is imported into cells, it will be reduced to cysteine in a short time [35]. Mutations of key genes in cysteine metabolism pathways were correlated with abnormal cysteine concentration [30], which might be correlated with disease development and progression. For example, mice with a knockout in *EAAC* gene, which was known to be cysteine transporter in central neuron system, would develop a mental retardation syndrome similar to Parkinson disease [36]. Similarly, silencing of cystathionine γ-lyase led to lower cysteine concentration and mediated neuro-degeneration in Huntington's disease mouse model [37].

Because neurodegenerative diseases such as Parkinson disease had a low glutathione concentration in the plasma [38], it was difficult to track disease progression using glutathione sensor. In this case, the intracellular level of cysteine should provide information about signal transduction, nutrition availability and anti-oxidant effect during disease progression.

1.1.3 Methods to determine cellular cysteine concentration

Several methods have been developed for the measurement of cysteine concentration, including high performance liquid chromatography (HPLC) [39], capillary gas chromatography-mass spectrometry [40], capillary electrophoresis [41], colorimetric and fluorescence detection [42]. Most chromatography methods

are highly dependent on equipment and hard to be processed in high throughput within a short timeframe [43]. Thus great efforts have been contributed to the development of selective fluorescence probe for cysteine detection inside cells recently [44, 45].

Tanaka et al. reported a fluorescence turn-on cysteine sensor in 2004 [44]. However, the dye had a maximum emission wavelength at 380 nm, which was not within the visible region and could not be tracked by a regular microscope. Another turn-on cysteine/homocysteine sensor was developed by Zhang et al. in 2007 to detect the intracellular cysteine/homocysteine concentration [45], nevertheless the feasibility of the usage of this dye in live cells staining was questionable.

In 2014, Dr. Honglin Li from our lab published a paper to correct the structure of cysteine sensor that was reported in Zhang et al.'s paper [46]. Based on this corrected core structure, we synthesized a family of new probes with turn-on fluorescence property when reacting with cysteine and homocysteine. In this chapter, the application of these dyes to detect cysteine/homocysteine in live cells was presented. Since intracellular cysteine concentration was significantly higher than homocysteine [9, 47], we hypothesized that the fluorescent signals were predominantly from the reaction with cysteine.

Cysteine/homocysteine staining protocols in live cells provided information about cysteine/homocysteine concentration and localization. We were really interested in the study of breast cancer, thus our methods were applied to determine cysteine concentration in breast cancer cell lines.

1.1.4 HER2 and breast cancer

Breast cancer is one of the most common types of cancer and a major cause of cancer-related deaths in women [48]. Certain factors have been reported to influence prognosis of breast cancer, including size, histological type and grade of the primary tumor, lymph node involvement, status of estrogen and progesterone receptors and other biomarkers [49]. Of those markers, the human epidermal growth factor receptor 2 (HER2) is one of the well-studied ones. Overexpression and amplification of *HER2* have been shown to be correlated with increased disease recurrence as well as poor prognosis [50, 51].

HER2 is a member of the epidermal growth factor receptor (ErbB) [52]. The ErbB family includes four members of plasma membrane bound receptor tyrosine kinases [53]. All four members contain an extracellular ligand binding domain, a transmembrane domain and an intracellular domain [54]. HER2 could form heterodimers with any other three receptors [55], of which dimerization would induce the auto-phosphorylation of tyrosine residues within the intracellular domains and initiated various downstream signal pathways leading to cell survival and cell proliferation (Fig. 1.3). Since overexpression of *HER2* occured in 15-30% of breast cancer, drugs targeting HER2 could be used as potential cancer treatment for HER2 positive patients [50, 51].

HER2 expression levels might be related to intracellular cysteine level. For example, it was reported that the expression of micro RNA miR-26b, which targeted a subunit of cystine transporter-SLC7A11 [56], was significantly reduced in HER2+ tumor specimens [57].

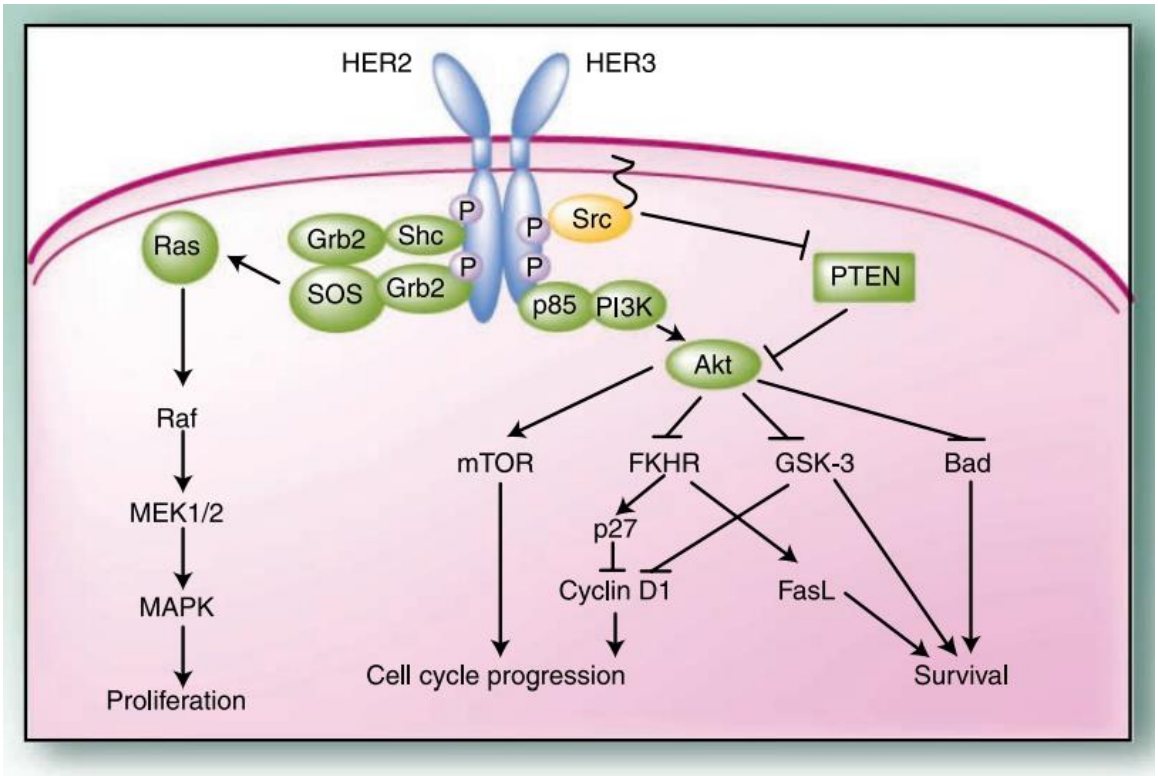


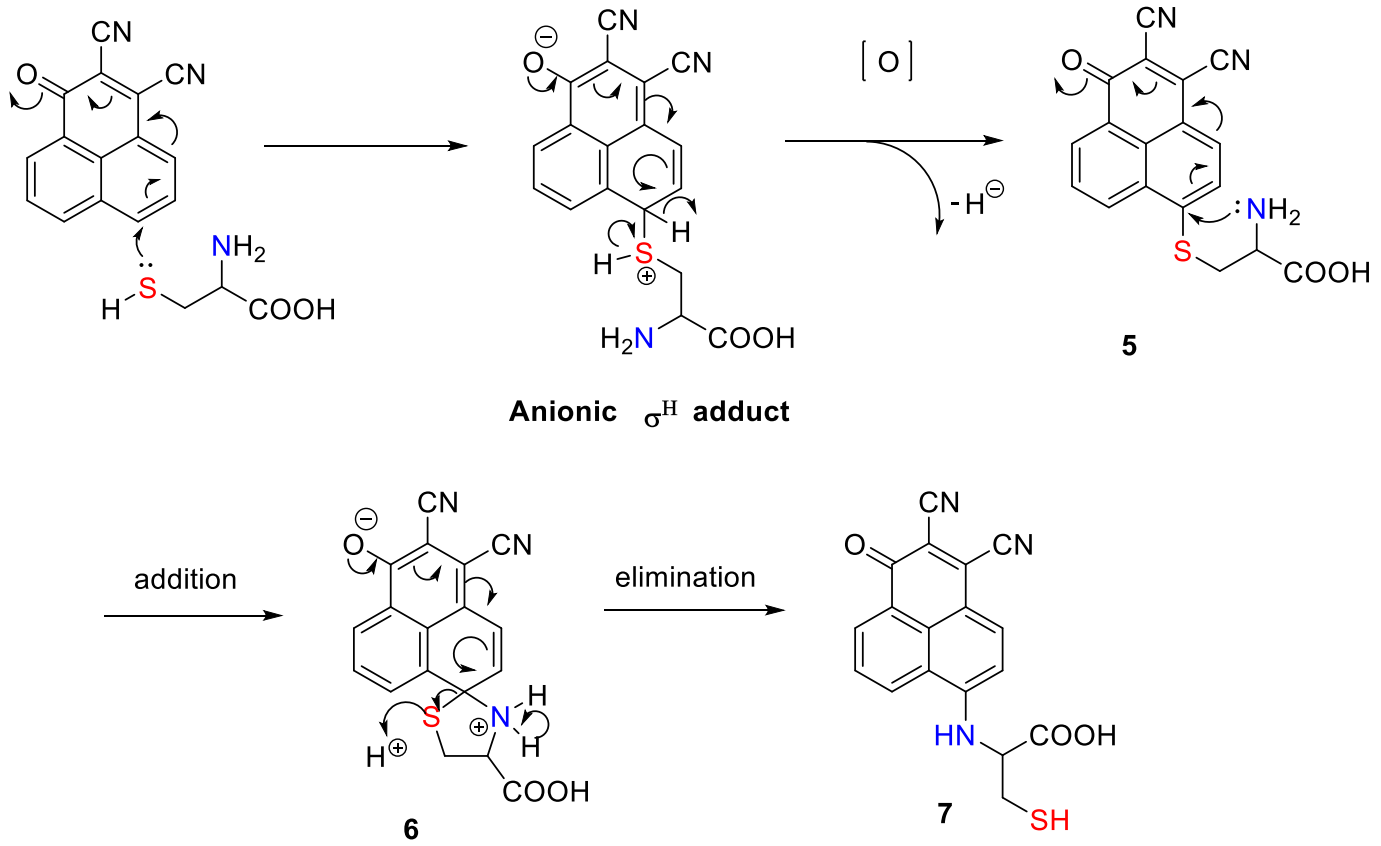
Figure 1.3. HER2 signaling pathway. Adapted from "N. U. Lin, et. al. 2007 Clin Cancer Res" with copy right permission.

To study the correlation between *HER2* expression and cysteine level, we used MCF10A cell lines which have been programmed to over express control *vector* or *HER2* (denoted as MCF20A *vector* and MCF10A *HER2*). MCF10A cell line has low expression of *HER2* [58] and does not introduce genetic interference from other oncogene mutations. Overexpression of *HER2* can consistently activate cell survival and proliferation pathways, which lead to higher tumorigenicity of this breast epithelial cell line [58]. Thus, MCF10A *vector* and MCF10A *HER2* cell lines could be used as models for normal breast cells and HER2+ breast cancer cells.

1.2 Results and Discussion

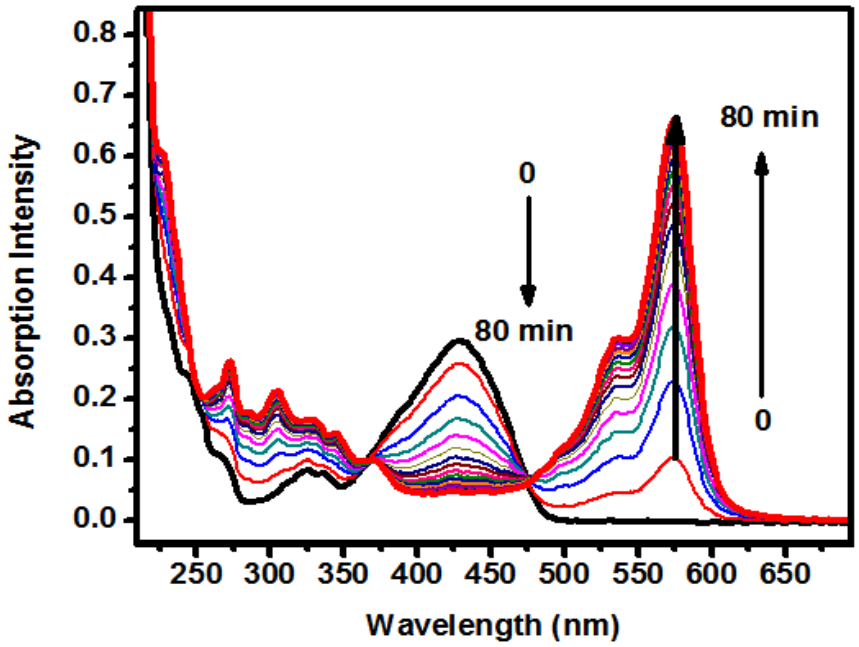
1.2.1 Design of fluorescence turn-on sensors

As showed in Fig. 1.2, a series of fluorescent probes with 1-oxo-1*H*-phenalene-2,3-dicarbonitrile core structure have been synthesized in our lab. A unique sensing mechanism was proposed for this type of sensors (Scheme 1.1): the intramolecular base promoted oxidative S_NAr^H (nucleophilic substitutions of aromatic hydrogen) reaction combined with an intramolecular displacement S_NAr^{ipso} (nucleophilic aromatic substitution) of thiol group by amino groups (Scheme 1.1). All three sensors react quickly with thiols in the presence of base and much slower under neutral condition (unpublished data). The absorption spectrum and fluorescence spectrum of reaction were measured on Varian Cary Eclipse Fluorescence Spectrophotometer. The maximum absorption peak of sensor **1** was centered at 430 nm (Fig. 1.4.a). Upon reaction with simple thiol-containing compounds, the absorption peak shifted to 512 nm, while the peak further shifted to 574 nm



Scheme 1.1. Proposed mechanism of the oxidative S_NAr^H and the following S_NAr^{ipso} reaction between sensor **1** and cysteine

a)



b)

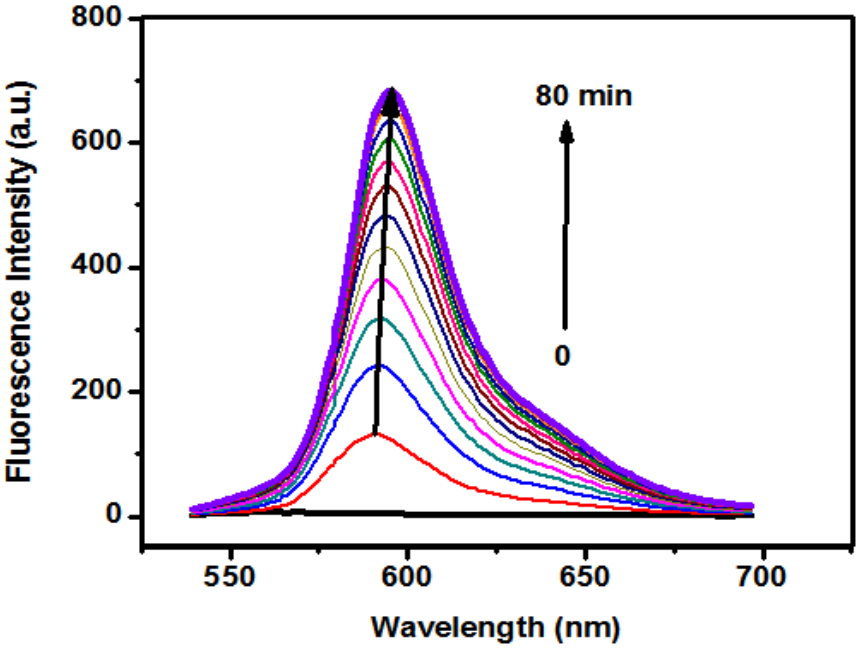


Figure 1.4. Absorption (a) and fluorescence (b) spectra changes of 20 μM sensor **1** in CH_3CN from 0 to 80 min upon addition of 100 μM cysteamine.

when reacted with amino acid. The fluorescence spectrum of reaction products showed peaks centered at 582 nm for simple thiol and 592 nm for amino acid. Due to the lack of neighboring amino group, such rearrangement reaction would not take place when the probes reacted with GSH. Instead, the probes could only react slowly with GSH through a direct thiol-substitution reaction. Because sensors **1-3** nearly showed no fluorescence, their turn-on property (Fig 1.4.b) and peak shift of fluorescence allowed them to be good candidates for intracellular cysteine/homocysteine sensing.

1.2.2 Sensing specificity for cysteine

To test the specificity of these turn-on sensors, different substrates including amino acids, thiol containing molecules (GSH, N-acetyl cysteine), dNTP and protein (bovine serum albumin) were incubated with sensors for 2 hours to allow reactions processing in phosphate buffered saline (PBS) (Fig. 1.5. a). The fluorescence was recorded on MDS SpectraMax M2 microplate reader in 96-well microtiter plates. The folds changes of fluorescence: $\mathbf{(I-I_0)/I_0}$ (**I**: fluorescence intensity of indicated solution; $\mathbf{I_0}$: fluorescence intensity of background.) did not change for more than 1 for most substrates (at the range of -0.45 ~ 0.42). This indicated that all three sensors had comparable low backgrounds in fluorescence when were incubated with those substrates.

Conversely, the fluorescent intensity of sensor **1** increased for 43.8 ± 6.6 times when incubated with cysteine. Around 3.7 ± 0.4 and 2.0 ± 0.2 fold increase in fluorescence was observed for sensor **2** and **3**. The fluorescence "turn-on" fold was also observed when treating with homocysteine, but it was much less profound, i.e. 19.1 ± 0.5 folds increase for sensor **1**, 1.8 ± 0.9 folds for sensor **2** and 0.75 ± 0.06 folds for sensor **3**, respectively.

Although the sensitivity of sensor **1** was much higher than sensor **2** and **3** on the detection of cysteine/homocysteine, the standard deviation was also much bigger for sensor **1** (Fig. 1.5).

The intracellular concentration of GSH was very high and around 1-10 mM [14, 15], while the intracellular concentration of cysteine was relatively low and around 50-200 μM [2, 47]. Although we had confirmed that the fluorescent signal was comparable to background when incubating the sensors with 50 μM GSH, it might not be the same when GSH concentration was within the intracellular range, which might be 200 fold higher. To identify the potential interference from high concentration GSH, we incubated those sensors with either 200 μM cysteine or 5 mM GSH for 2 hours. The fluorescence signals for cysteine/GSH/background were: sensor **1** (267.5/19.6/4.2), sensor **2** (144.9/26.3/19.1) and sensor **3** (101.3/244.3/34), respectively (Fig. 1.5.b). This result confirmed the feasibility of using these sensors for intracellular cysteine detection.

1.2.3 Quantitation of cysteine in vitro

Before using our designed sensors to quantify cysteine, the relationship between cysteine concentration and fluorescence intensity needed to be demonstrated. Titrations of cysteine concentration against fluorescence intensity were performed for all three sensors in DMEM and the standard curves were generated (Fig. 1.6.a). Three calibrations were quadratic with coefficients of determination (R^2) value around 0.99. This reliable relation between fluorescent intensity and cysteine concentration supports sensor **1**, **2** and **3**'s possibility to

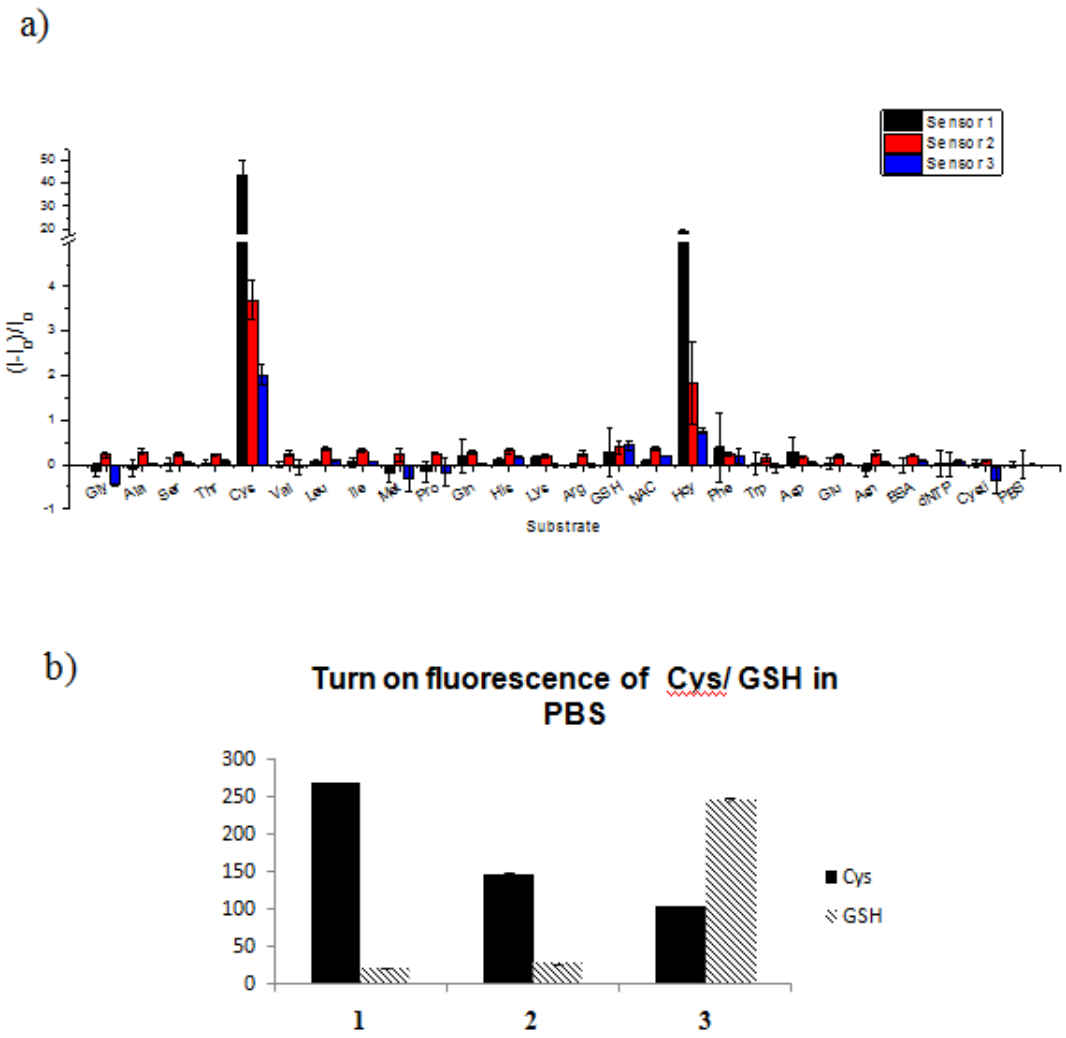


Figure 1.5. Fluorescence selectivity of cysteine sensor **1-3**. a) Fluorescence intensity on different substrate in PBS. All substrate: 50 μM. **Gly**: Glycine; **Ala**: Alanine; **Ser**: Serine; **Thr**: Threonine; **Cys**: Cysteine; **Val**: Valine; **Leu**: Leucine; **Ile**: Isoleucine; **Met**: Methionine; **Pro**: Proline; **Gln**: Glutamine; **His**: Histidine; **Lys**: Lysine; **Arg**: Arginine; **GSH**: Glutathione; **NAC**: N-Acetylcysteine; **Hcys**: Homocysteine; **Phe**: Phenylalanine; **Trp**: Tryptophan; **Asp**: Aspartic Acid; **Glu**: Glutamic Acid; **Asn**: Asparagine; **BSA**: Bovine Serum Albumin; **dNTP**: deoxynucleotide; **Cysti**: Cystine. b) Selectivity of sensors to intracellular level cysteine (200 μM) over GSH (5 mM). Substrates were incubated with indicated sensors for 2 h and fluorescence was read on plate reading. Sensor concentration for a) and b): **1** (10 μM) **2** (10 μM), **3** (50 μM).

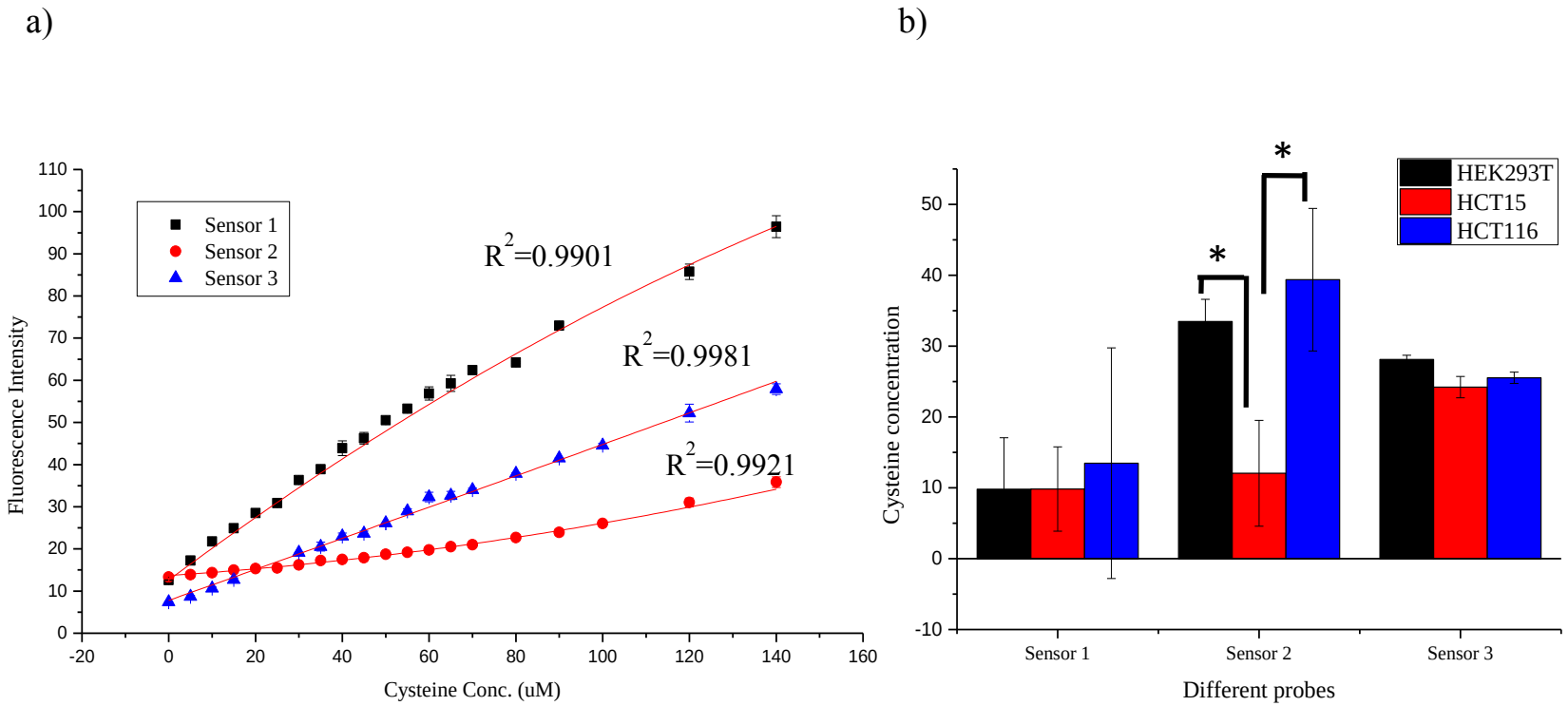


Figure 1.6. Quantification of cysteine *in vitro*. a) Standard curve of sensor **1** (black square), **2** (red circle) and **3** (blue triangle). The polynomial fit of each sensor was showed in red line. b) Quantification of cysteine in cell culture medium. Sensor concentration for a) and b): **1** (10 μM) **2** (10 μM), **3** (50 μM). * indicate $p < 0.05$.

quantify cysteine concentration.

To evaluate the potential application of cysteine sensors *in vitro*, the concentration of cysteine in cancer cell conditional medium was measured using this method. Conditional medium from HEK293T, HCT15 and HCT116 were collected as described in the experimental section. The medium was incubated with sensors for 0.5 hour and the fluorescence was recorded. Cysteine concentrations in the medium were calculated from the calibration. As showed in Fig.1.6.b, the cysteine concentrations in all three cell culture mediums were around 10 μM measured by sensor **1** and around 25 μM by sensor **3**. The reading of concentration for three cell lines was different from sensor **2**, i.e. HEK293T and HCT116 medium were 33.5 ± 3.1 μM and 39.4 ± 10.0 μM while HCT15 was 12.1 ± 7.5 μM respectively. There was no cysteine but only cystine in prepared mediums for HEK293T, HCT116 and HCT15. While cells were cultured in those mediums, they could uptake cystine and converted it to cysteine, which were then released to the mediums and might be used by other cells. From this experiment we found all three cell types could uptake cystine and convert it to cysteine. The cysteine concentrations in cell conditional medium were around 40 μM and might be varied among cell lines.

1.2.4 Cell compatibility of cysteine sensors

After confirming the *in vitro* application of these sensors, the cell compatibility of those sensors for live cell imaging was investigated. Sensor **1** and **2** were used to stain the cells and confocal images were taken to track their fluorescence. Sensor **3** was not tested due to its nonspecific reaction with GSH of intracellular concentration.

As shown in Fig. 1.7, the staining of sensor **1** was much weaker and not evenly distributed. This might be attributed to the low solubility of sensor **1** in aqueous solutions. Sensor **2** showed a brighter fluorescence signal and the staining was evenly distributed among cells. Since Sensor **2** had better cell compatibility, it was used in the consequent live cell imaging studies.

1.2.5 Localization of cysteine/homocysteine in live cells

One advantage of our turn-on fluorescence sensors was the low background, by which the tracking of cysteine in cells was clearer and more reliable. Cysteine is distributed all over the cells with enrichment in certain organelles such as mitochondria and lysosomes for their specific functions inside cells [45, 59]. Sensor **2** had shown an uneven intracellular cell staining. There were regions that had higher fluorescence signals (Fig. 1.8.c and Fig. 1.8.g). We found that those regions with higher fluorescence signals were co-localized with both MitoTracker Green and LysoTracker Green (Life Science). Since intracellular homocysteine is negligible compared to cysteine, the majority of the fluorescence signal can be attributed to the reaction with cysteine. Although we have not performed further experiments to study the local cysteine concentration changes inside cells, it will be extremely interesting to know the cysteine concentration in specific localizations i.e. mitochondria and lysosome under different stimuli, especially for cancer cells. These cells are usually under high oxidative stress [30], which results in higher cysteine concentration in those localizations.

1.2.6 Semi-quantitation of intracellular cysteine in live cells

As shown in section *1.2.3*, the fluorescence intensity after incubating with sensor was proportional to cysteine concentration. Because sensor **2** had excellent cell compatibility, it was possible to use sensor **2** to quantify the intracellular concentration of cysteine. The intracellular cysteine concentrations are known to be 150 μM [47]. Certain cell lines, such as Hepg2/C3A, exhibits 3-6 fold changes in cysteine concentration when cultured with different concentrations of cysteine in medium [11, 60]. We pretreated Hepg2/C3A and HCT116 for 24 hours with 0 or 400 μM of cysteine in the full culture medium before the cells were washed and incubated with sensor **2** for 0.5 hour. Confocal images and flow cytometry were performed to track the intracellular cysteine concentration. Images from confocal microscope (Fig. 1.9.a) showed much higher fluorescence intensity in cells treated with 400 μM cysteine, which was further confirmed by flow cytometry (Fig. 1.9.b). Approximately 1.7 folds increase in fluorescence signal (Fig. 1.9.c) was observed in both cell lines (results of HCT116 were showed in Fig. 1.10) for cells incubated with 400 μM cysteine. These staining results confirmed sensor **2**'s function on specifically indicating intracellular cysteine in live cells.

To further study the cysteine concentration in cancer cells, we chose two MCF10A derived cell lines: MCF10A *vector* and MCF10A *HER2*. Cells were incubated with sensor **2** for 0.5 hour and the fluorescence was tracked by confocal imaging and flow cytometry. The fluorescence staining (Fig. 1.11) was higher for MCF10A *HER2* cells and flow cytometry results confirmed MCF10A *HER2* had higher fluorescence intensity (7279) than MCF10A *vector* (4380). It indicated that breast cancer cells had higher cysteine concentration compared to normal cell lines.

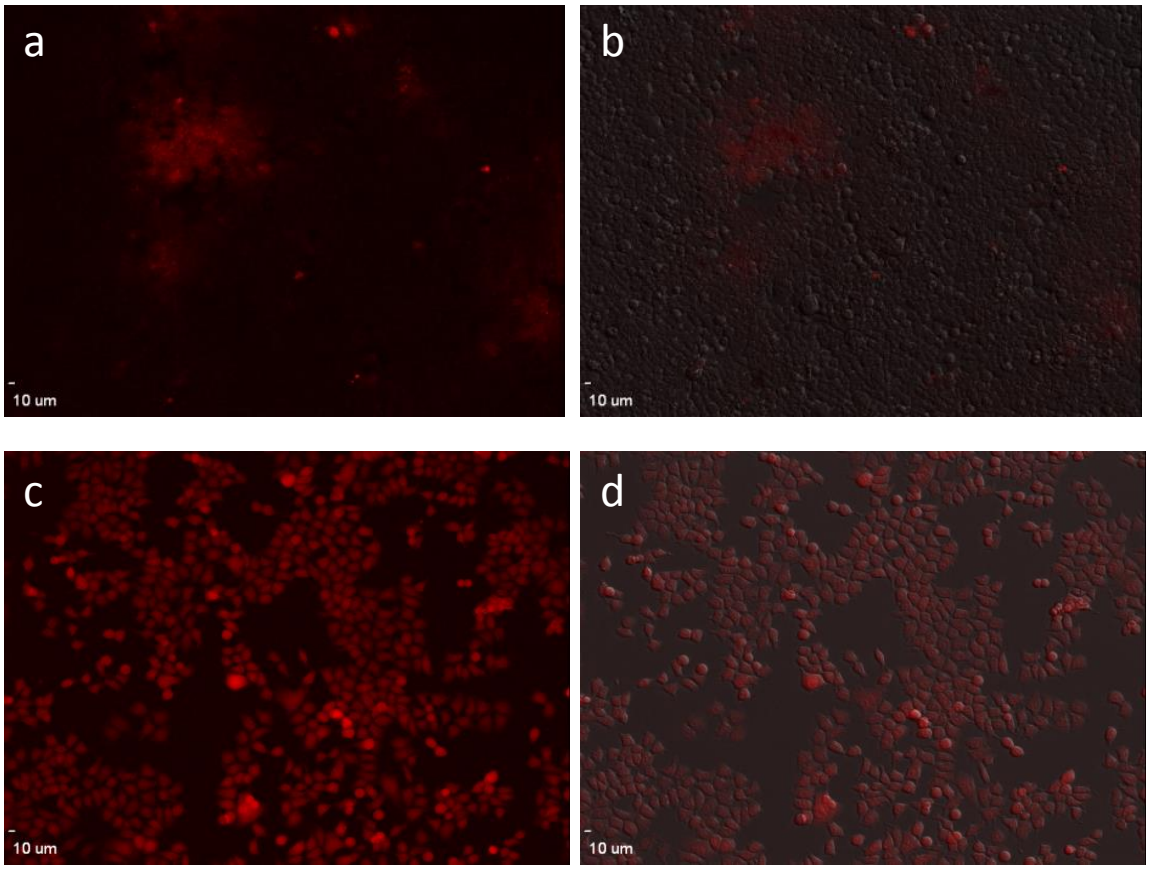


Figure 1.7. Fluorescent images (a, c) and overlay image with bright field (b, d) of HCT116 cell lines stained with 50 μM sensor **1** (upper level) and sensor **2** (lower level) for 0.5 h.

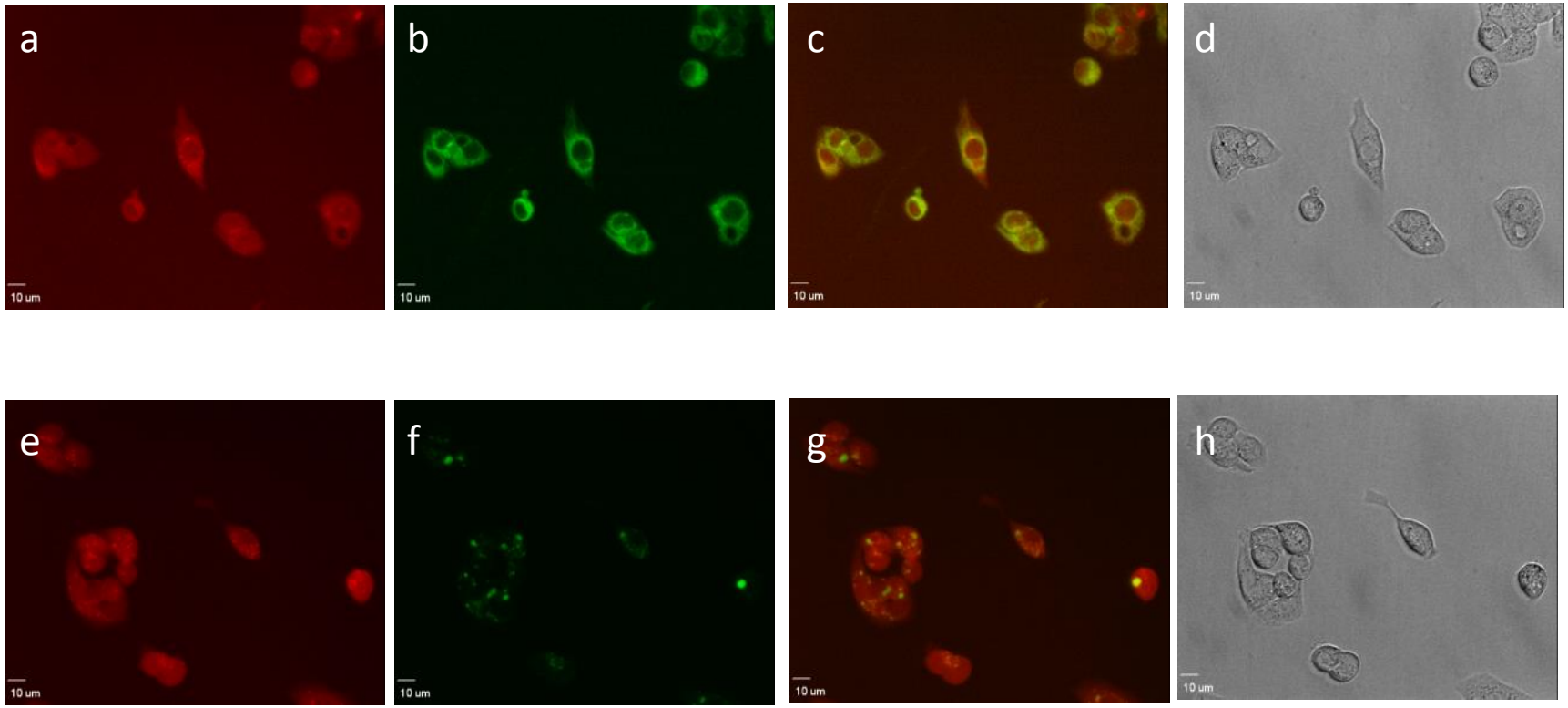


Figure 1.8. Fluorescence (a-c, d-g) and bright field images (d, h) of live Hepg2/C3A. Upper: Hepg2/C3A was co-stained with 50 μM sensor **2** (red) and MitoTracker (green) for 0.5 h. Bottom: Hepg2/C3A was co-stained with 50 μM sensor **2** (red) and LisoTracker (green) for 0.5 h.

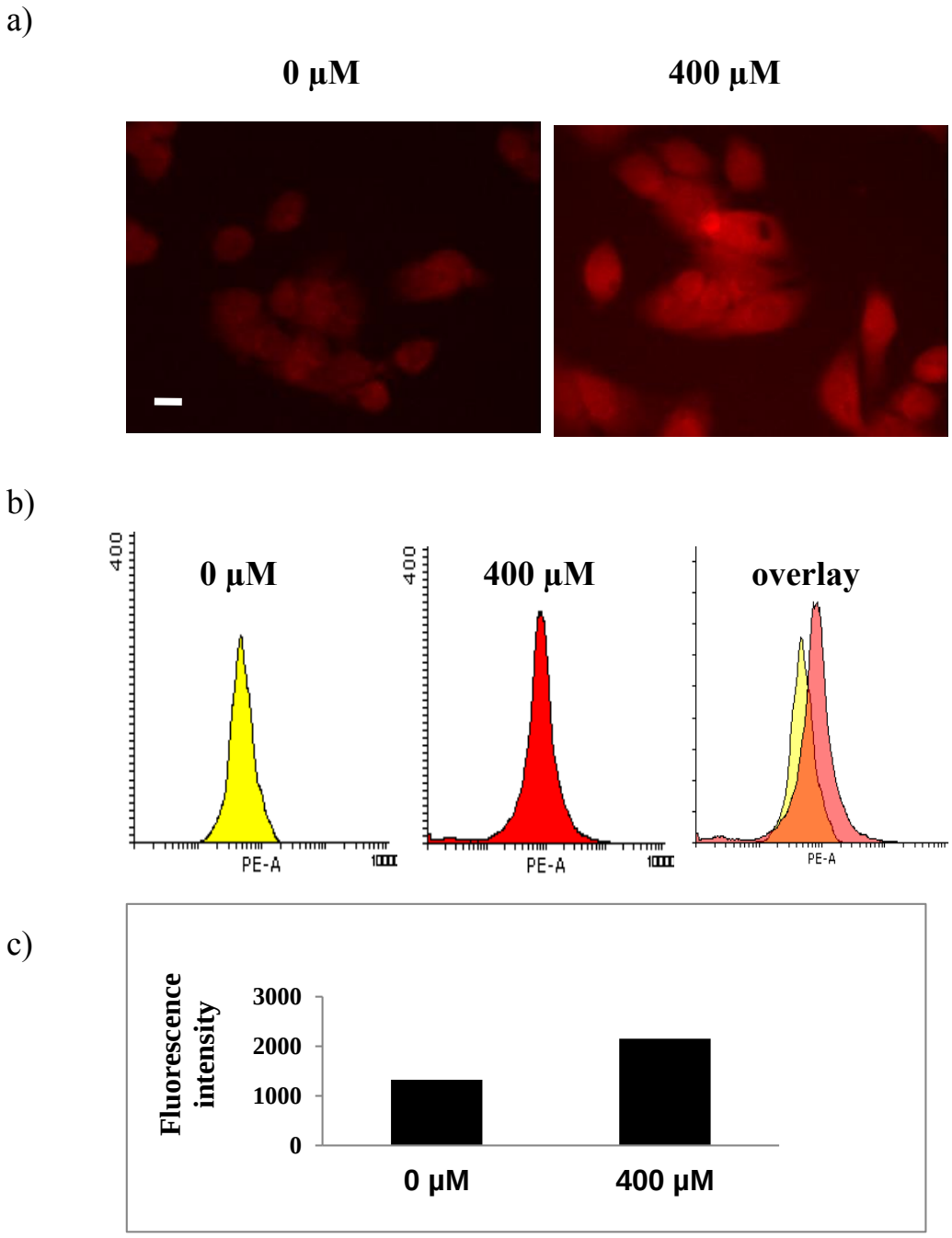
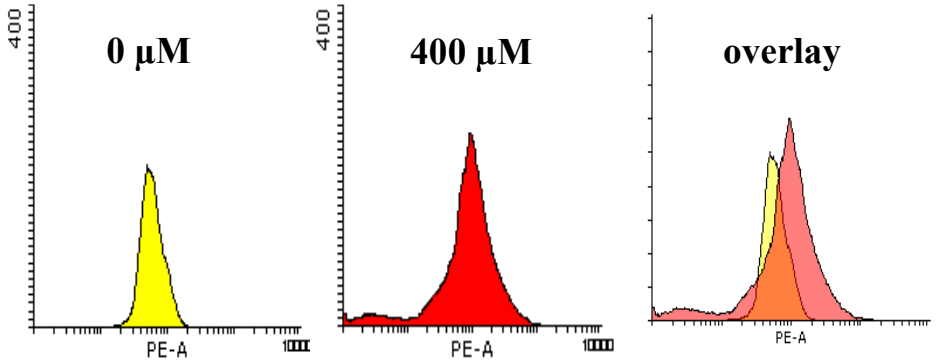


Figure 1.9. Sensor **2** indicated intracellular cysteine concentration in live cells. Hepg2/C3A were pretreated with different concentration of cysteine for 24 h and then stained with 50 µM sensor **2** for 0.5 h. a) Confocal images of Hepg2/C3A pretreated with 0 or 400 µM Cysteine then stained with sensor **2** (scale bar: 10 µm). b) Flow cytometry of the Hepg2/C3A pretreated with 0 or 400 µM cysteine (yellow: 0 µM; red: 400 µM). c) Bar graphs showing the fluorescence quantification of the Hepg2/C3A treated with 0 or 400 µM cysteine.

a)



b)

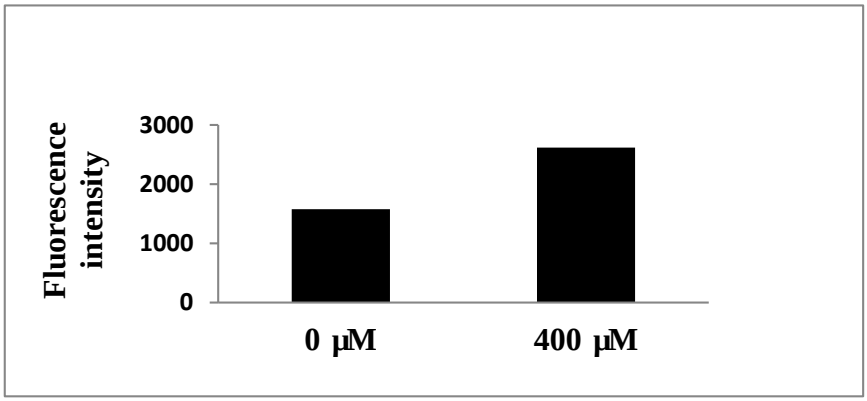


Figure 1.10. Sensor **2** indicated intracellular cysteine concentration in live cells. HCT116 were pretreated with 0 or 400 µM cysteine cysteine for 24 h and then stained with 50 µM sensor **2** for 0.5 h. a) Flow cytometry of the HCT116 pretreated with 0 or 400 µM cysteine then stained with sensor **2** (yellow: 0 µM; red: 400 µM). b) Bar graphs showing the fluorescence quantification of the Hepg2/C3A treated with 0 or 400 µM cysteine.

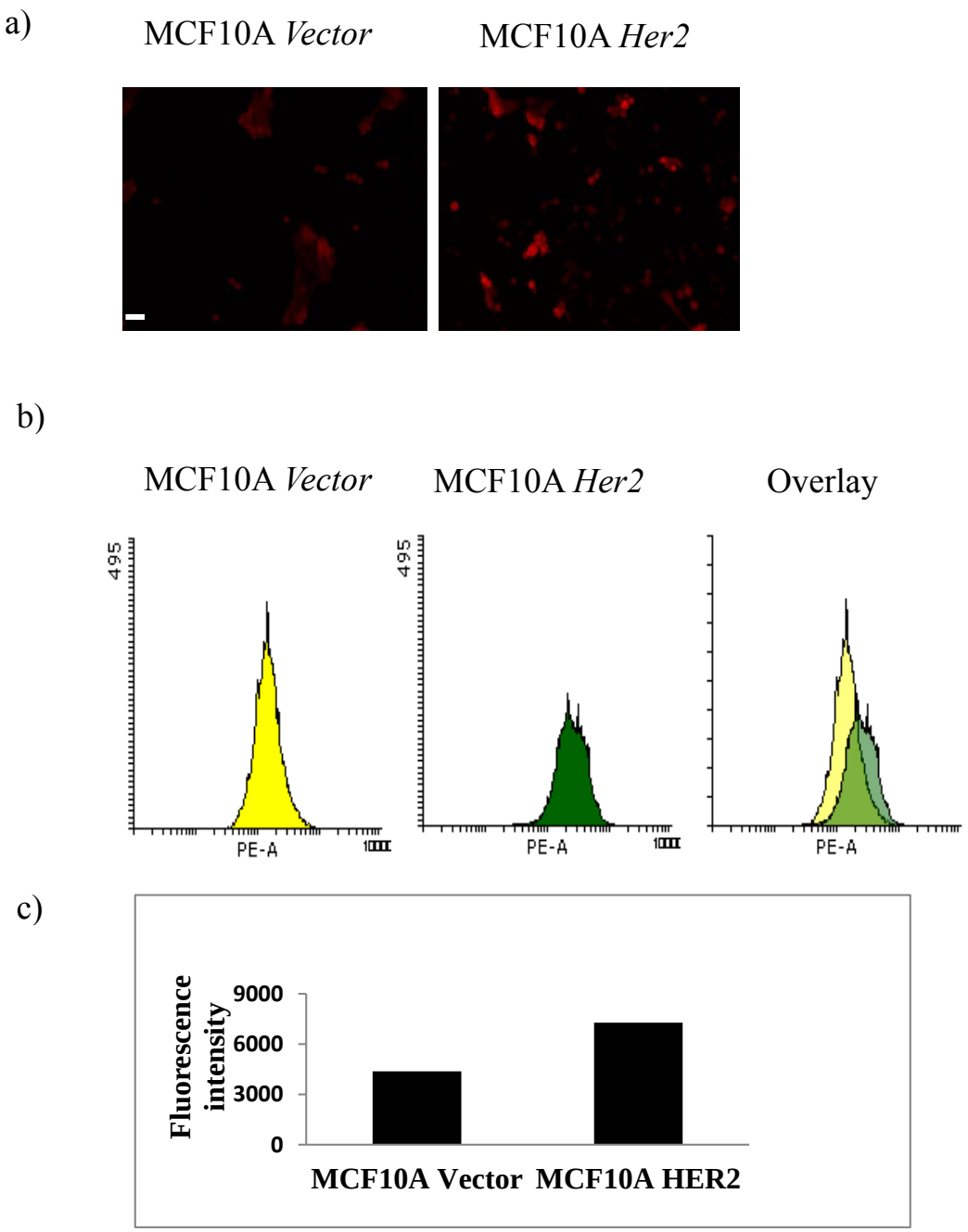


Figure 1.11. Cysteine concentration in MCF10A *Vector* and MCF10 *HER2* cell lines. a) Confocal images of MCF10A *Vector* and MCF10 *HER2* cell lines stained with sensor **2** (scale bar: 50 μm). b) Flow cytometry of the MCF10A *Vector* and MCF10 *HER2* cell lines stained with sensor **2** (yellow: MCF10A *Vector*; green: MCF10 *HER2*). c) Bar graphs showing the fluorescence quantification of the flow cytometry.

1.3 Conclusion

Cysteine, homocysteine and GSH are major thiol-containing biomolecules in cells. They play important roles in maintaining normal cell state. Homocysteine is the precursor to synthesis cysteine and cysteine works as precursor for GSH synthesis. Interferences of thiol metabolism either by mutation of key enzymes or transporters are proved to be correlated with many diseases, including cardiovascular diseases, cancer and neurodegenerative diseases. Current methods to measure thiol-containing molecules include chromatography and fluorescent sensors. Protocols that can specifically assess cysteine, homocysteine and GSH concentration in live cell are not well developed.

Our lab has synthesized a series of sensors that can selectively detect cysteine/homocysteine. The reaction was proved to be an intramolecular base promoted oxidative S_NAr^H (nucleophilic substitutions of aromatic hydrogen) reaction combined with intramolecular displacement (S_NAr^{ipso}) of thiol group by amino group. Three probes (**1**-**3**) all showed good selectivity towards cysteine/ homocysteine. Standard curves of three sensors from titration of cysteine fitted polynomial regression with R^2 over 0.99.

Sensor **1** and sensor **2** both could react with intracellular concentration of cysteine but not GSH. Thus, these two sensors were used for *in vivo* imaging of cysteine/ homocysteine. Because intracellular homocysteine concentration was significantly lower compared to cysteine concentration, fluorescence was mainly from the reaction with cysteine.

Both dyes could enter the cells (perhaps by using diffusion mechanism). Sensor **2** was chosen for its better water solubility and good cell staining results. Furthermore, we

demonstrated that sensor **2** could be used to semi-quantify the intracellular cysteine quantification using a hepatocellular carcinoma cell model. The live cell imaging not just provided the localization of cysteine but also illustrated a semi-quantification of cysteine in live cells.

We also confirmed that cysteine concentration was higher in HER2+ breast cancer cell line: MCF10A *HER2* compared to the control cell line: MCF10A *Vector*. With sensors **1**-**3**, we improved the study of changes of cysteine under different conditions or treatments, by which the changes of cysteine levels could be measured in the progressions of different diseases such as cancer, Parkinson's disease and Alzheimer disease.

1.4 Experimental Section

1.4.1 Sensing selectivity test of three sensors

Different substrates (50 μM each for final concentration) was added to 200 μL either PBS or DMEM with indicated concentration of sensors (sensors **1**: 10 μM, sensor **2**: 10 μM and sensor **3**: 50 μM) in black 96 well plate with clear bottom (Greiner 655906). Solutions were incubated for 2 hours at room temperature. The fluorescence was recorded on a MDS SpectraMax M2 Microplate Reader with excitation at 560 nm and emission at 600 nm.

1.4.2 Cysteine standard Curve

Different concentrations of cysteine were added to 200 μL DMEM with indicated concentration of sensors (sensors **1**: 10 μM, sensor **2**: 10 μM and sensor **3**:

50 μM) in black 96 well plate with clear bottom (Greiner 655906). Solutions were incubated for 2 hours at room temperature. The fluorescence was read on a MDS SpectraMax M2 Microplate Reader with excitation at 560 nm and emission at 600 nm.

1.4.3 Cell Maintenance

The Hepg2/C3A cell line was purchased from the American Type Culture Collection (Cat No: HB/8065). Cells were maintained in sulfur amino acid-free DMEM supplement with 10% fetal bovine serum, 2 mM glutamine, 1 mM sodium pyruvate, 0.1 mM L-methionine, 0.3 mM cysteine and 1 × MEM nonessential amino acid solution. The HCT116, MCF10A *vector* and MCF10A *HER2* cell lines were from Prof. Hexin Chen (University of South Carolina, Department of Biology Science). The HCT15 cell line was from Prof. Franklin G. Berger (University of South Carolina, Center for Colon Cancer Research). The HEK293T cell line was from Dr. Daping Fan (University of South Carolina, Department of Biology and Anatomy). HCT116, HEK293T cells were maintained in DMEM-high glucose medium with 10% fetal bovine serum. HCT15 cells were maintained in RPMI1640 medium with 10% fetal bovine serum (FBS). The MCF10A *Vector* and MCF10A *Her2* were maintained in DMEM:F12(1:1) medium with 5% horse serum, 10 μg/mL insulin, 20 ng/mL EGF, 100 ng/mL cholera toxin, 0.5 μg/mL hydrocortisone. Cells were passed every 2-3 days on a (1:8) ratio. All cells were in a water-saturated atmosphere of 5% CO_2 and 95% air at 37 ℃.

1.4.4 Cell culture medium collection

To determine the cysteine concentration in different conditional medium, 1×10^6 cells/well HEK293T, HCT15 and HCT116 were seeded in 6-well plates and allowed to adhere and grow overnight. The medium was then switched to DMEM medium without FBS and cultured for another 24 hours. The medium was collected, clarified by centrifugation and used immediately.

1.4.5 Confocal imaging

Confocal fluorescence imaging was performed with Olympus X81 fluorescence microscopy. Cells were incubated with sensors and (+/-)-Verapamil Hydrochloride (MP Biomedicals, Cat. No: 195545) for 0.5 hour at 37 ℃. Fluorescence imaging was performed after washing cells three times with PBS buffer.

1.4.6 Flow cytometry

Cellular cysteine levels were quantified by flow cytometric analysis. 2×10^5 cells/well HepG2/C3A cells were seed on 6-well plates two days before flow cytometry. The condition cell mediums with different cysteine concentration were changed one day before flow cytometry. Cells were trypsinized and 1×10^6 cells were collected for each sample as indicated. After washing two times with Assay Buffer (PBS with 0.5% FBS and (+/-)-Verapamil Hydrochloride), cells were incubated with 50 μM sensor **2** for 0.5 hour at 37 ℃. The samples were washed once with Assay Buffer and re-suspended in 0.5 mL Assay Buffer. Stained cells were analyzed by BD LSRII flow cytometer.

CHAPTER 2 Cell Toxicity of DCP Family Probes

2.1 Introduction

2.1.1 DCP family may have potential cell toxicity towards cancer cells

The DCP family of probes have been reported by Dr. Xuhong Qian's group and his coworkers as a potential inhibitor for Fibroblast Growth Factor [61]. Using enzyme-linked immunosorbent assay (ELISA), three compounds were found to be active in inhibiting the activation of the Fibroblast Growth Factor Receptor 1 (FGFR1). Cell studies also confirmed their function in anti-proliferation of tumor cells [61]. More papers on the DCP derivatives were published in 2011 and 2013 [62, 63]. Certain DCP analogues could specifically inhibit the activation of the B-Cell Lymphoma 2 (Bcl-2) and Myeloid Cell Leukemia Sequence 1 (Mcl-1), thus induced apoptosis in tumor cells [62, 63].

Recently, we found the structures assigned in both Xuhong Qian and Zhichao Zhang groups' papers were not correct [46]. Dr. Honglin Li from our group confirmed the core structures of the DCP family compounds as 1-oxo-1*H*-phenalene-2,3-dicarbonitrile. To further investigate the potential roles as antitumor agents for these molecules, a series of DCP derivatives were synthesized. The function of those analogues as antitumor drugs was discussed in this chapter.

We used MCF10A cell lines that had been programmed to over express control vector and *HER2* as models. The DCP family molecules were used to treat those cell

lines and the cell viability was tested. The related signaling pathways were also investigated in order to discover the target of those molecules.

2.2 Results and Discussion

2.2.1 Structure of DCP derivatives

The DCP family probes were shown in Fig. 2.1. They share the same 1-oxo-1*H*-phenalene-2,3-dicarbonitrile core structure, with different side chains. Previous studies from Dr. Qian's lab showed molecules with thiol containing side chains at position 3 could inhibit tyrosine kinase activation, including FGFR1, FGFR3 and VEGFR1. As we known, receptor tyrosine kinases were highly conserved in their intracellular C-terminal, which was responsible for the kinase activity. We therefore hypothesized that DCP family derivatives, which may be an inhibitor of kinase activity, might also inhibit HER2 activation. To test this hypothesis, our lab modified the core structure at position 6 and position 9 of DCP with thiol-containing side chain, azide-containing side chain and nitro-containing side chain for screening the impact of derivatization of DCP on HER2 inhibition.

2.2.2 Cytotoxicity of DCP and derivatives

The treatment for HER2 positive breast cancer was by a monoclonal antibody (Trastuzumab) [64]. Other small molecules as HER2 tyrosine inhibitors also underwent active clinical trials for their advantages in synthesis and treatment [64]. To screen for HER2 specific drugs, we chose MCF10A cell lines which were consistently over-expressing either control vector or *HER2,* namely MCF10A *Vector* and MCF10A *HER*2 as shown in Chapter 1. The MCF10A *control* could be

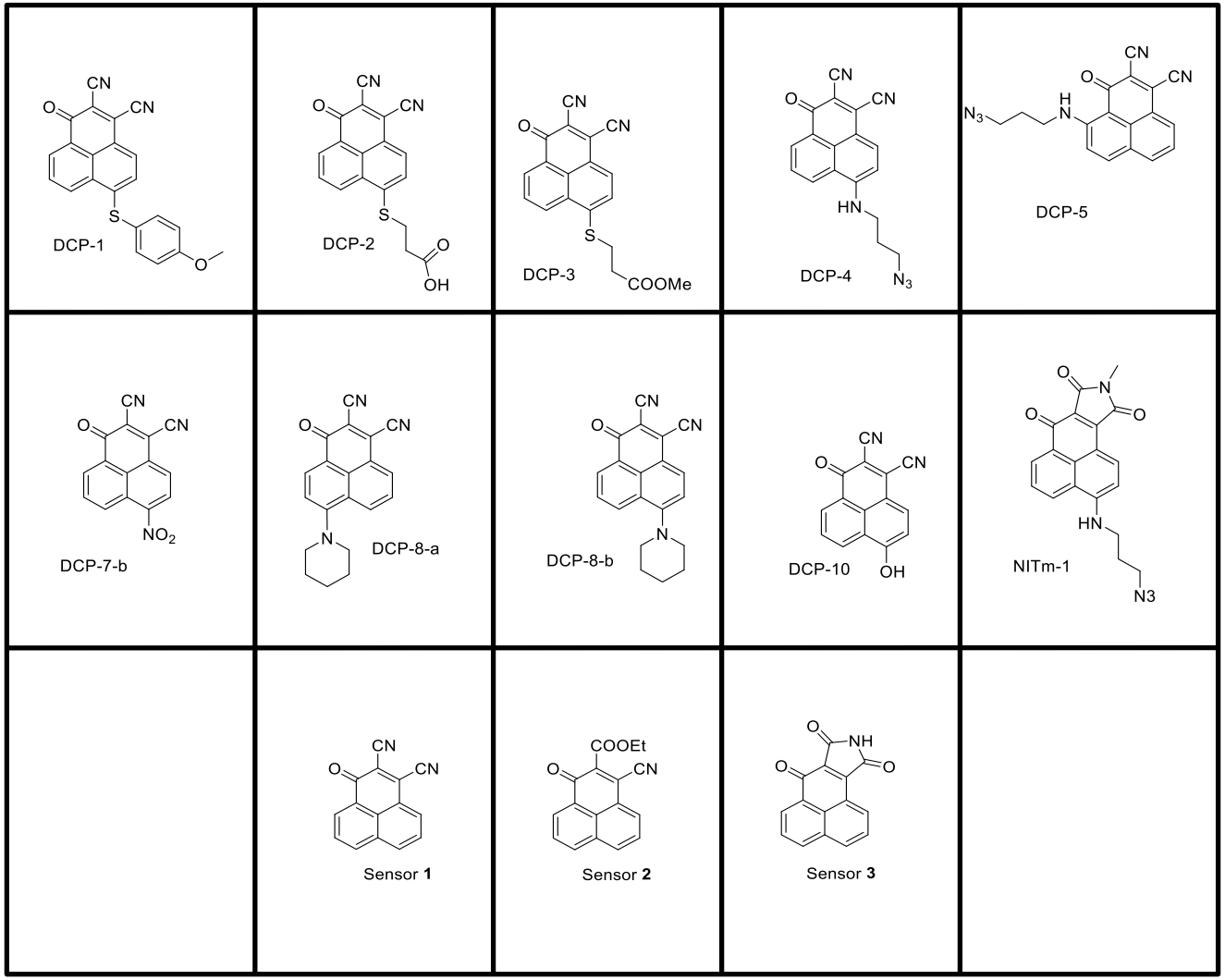


Figure 2.1. Structure of DCP and its derivatives tested in our study.

considered as normal breast tissue model while the MCF10A *HER2* cell line could be considered as HER2+ breast cancer model. We screened all molecules at 1×10^{-6} M or 1×10^{-5} M for 24 hours using cell viability assay.

The results of the cell viability assay were showed in Fig. 2.2. Star in the figure indicated a significant difference in viability between MCF10A *control* and *HER2* cell lines by student t test with 95% confidence interval. At 1 μM, all DCP probes did not show significant toxicity. None of them could kill more than 50% cells. At 10 μM, several members like DCP8b, sensor **1** and sensor **3** showed over 50% toxicity towards MCF10A *HER2* cell line. DCP8b killed 66.3% cells while sensor **1** killed 94.8% cells and sensor **3** killed 85.3% cells. The student t test of viability between MCF10A *vector* and MCF10A *HER*2 groups indicated sensor **3** preferably killed MCF10A HER2 cell lines. Therefore, sensor **3** might be a good candidate for HER2 breast cancer treatment. Interestingly, DCP8a, which was only different from DCP8b by the position of side chain, was much less toxic than DCP8b.

According to NCI anti-cancer drug screening procedures [65, 66] in cell lines, IC_{50} (the concentration of treatment to kill 50% cells) should be less than 1×10^{-4} M to be considered for further studies. Thus, DCP candidates that could kill more than 50% MCF10A *HER2* cells were chosen (50% cell viability was showed in red dash line) for later downstream signaling pathway studies.

2.2.3 Signaling pathway targeted by drug candidates

To study the possible signaling pathways affected by drug candidates, Western-blots were performed to study the major cell signaling pathways in breast cancer cells. Cells were treated with DCP8a, DCP8b, sensor **1** and sensor **3** for 24 hours.

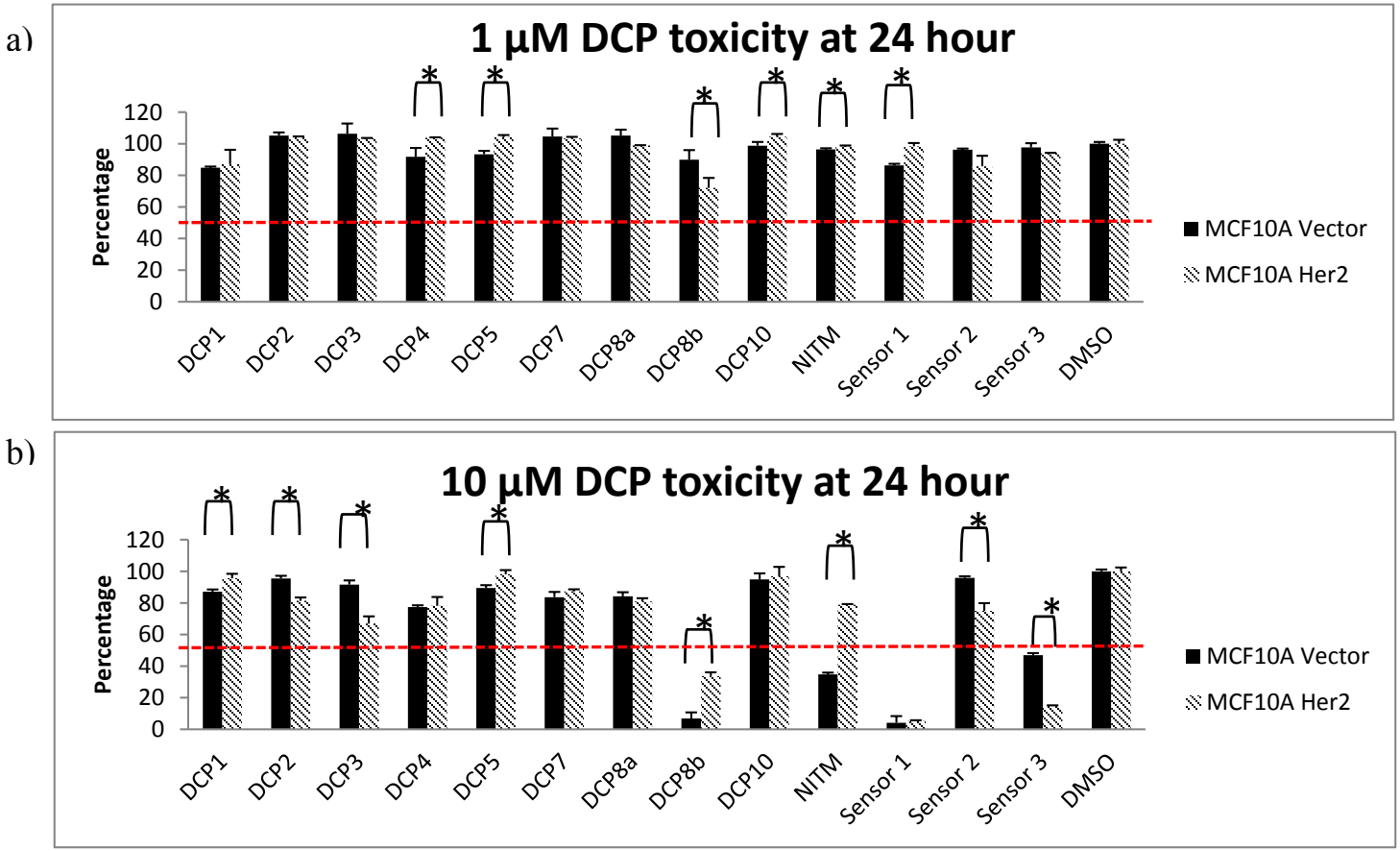


Figure 2.2. Toxicity of DCP family probes in different breast cell lines. a) The toxicity of 1 µM DCP family probes treatment. b) The toxicity of 10 µM DCP family probes treatment. Probes were incubated with cells for 24 h at indicated concentration. Star indicates significant difference in toxicity between MCF10A *vector* and MCF10A *HER2* cell lines by student t test with a 95 confidence interval. The red dash line is the 50% cell viability line.

Cells were lysed and proteins were collected. Activation of ERK (Extracellular signal Regulated Kinases), Akt (Protein Kinase B), Caspase-3 (cysteine-dependent aspartate-directed protease-3) and HER2 were probed by Western-blot with collaboration with Dr. Hexing Chen's group, which showed DCP8b, sensor **1** and sensor **3** led to less activation of HER2 (Fig. 2.3). This indicated that these molecules might interfere with HER2 activation. As comparison, DCP8a did not affect the activation of HER2 at all.

As shown in Chapter 1 (Fig. 1.3), there were mainly two pathways for the downstream activation of HER2. One was the ERK pathway that led to cell proliferation and the other was the Akt pathway that helped cell survival and protects from apoptosis. Thus inhibition of ERK might lead to abolishment of proliferation, while the inactivation of Akt might lead to the activation of Caspase-3. As Caspase-3 was a member of caspase family, this might lead to programmed cell death (apoptosis) [67]. The results from the down-stream signaling pathway (Fig. 2.4) showed that sensor **3** might achieve its specific toxicity towards *HER2* cell line by inactivation of ERK and activation of Caspase-3.

DCP8b seemed to induce the inactivation of ERK and slight activation of Caspase-3.This might explain the toxicity of DCP8b by inhibiting cell proliferation and activating apoptosis. Both DCP8a and DCP8b inactivated Akt, but DCP8a did not affect the other two pathways. The differences in toxicity and downstream signaling pathway might be caused by the different structures of DCP8a and DCP8b, or DCP8a and DCP8b might interact with different targets in the cells. It will be interesting to study the specific binding targets of each probe in the future. Sensor **1** did not affect Akt, ERK and Caspase-3 activation. The toxicity of this

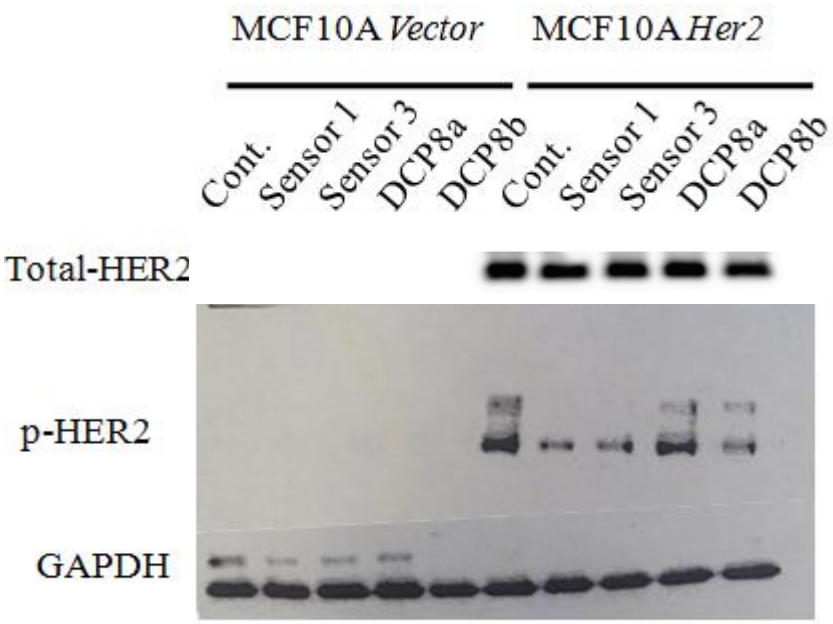


Figure 2.3. HER2 activation upon treating with DCP derivatives. Western-blot for total HER2 and p-HER2 showed that sensor **1**, sensor **3** and DCP8b can inhibit the activation of HER.

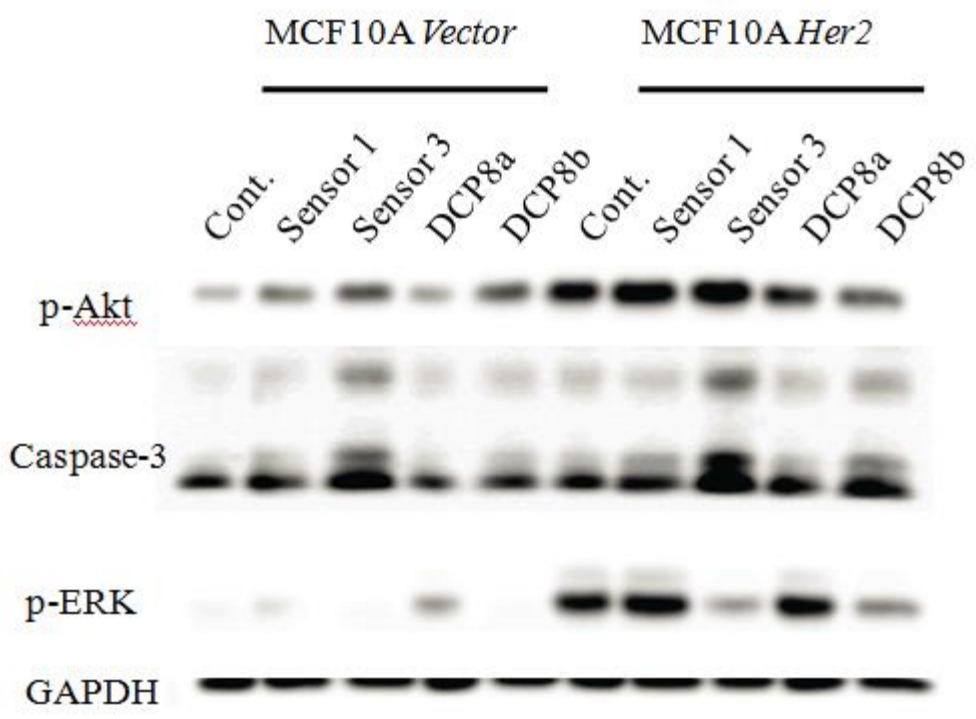


Figure 2.4. Cell signaling targeted by DCP candidates. The downstream signaling of HER2 were probes in Western-blot. p-ERK, p-Akt and Caspase-3 pathways were assayed to address the toxicity effect.

molecule to *Vector* and *HER2* cells lines might be attributed to some other pathways.

2.3 Conclusion

The core structure of DCP has been reported to be a good candidate as an anticancer drug. Our lab had successfully synthesized a series of DCP derivatives. To assess those probes' toxicity on cancer, we chose a specialized cell model for breast cancer. MCF10A either overexpressing control *vecto*r or *HER2* (i.e. MCF10A *Vector* and MCF10A *HER2*) were employed in our study to mimic normal breast tissue and HER2+ cancer. Using these two cell lines, we successfully screened the DCP derivatives and identified four candidates that have different toxic property: DCP8a, DCP8b, sensor **1** and sensor **3**. The results showed sensor **3** had higher toxic effect on MCF10A *HER2* cell lines.

To study the signaling pathway in those treatment groups, we collaborated with Dr. Hexing Chen's lab. Western-blots were performed to probe HER2 and its related pathways. It has been found that Sensor **3** treatment group significantly decreased HER2 activation. This indicated the specificity of toxic towards the HER2 cell line might be due to the inhibition of HER2. The ERK activation was also decreased, which indicated an important role of Sensor **3** on ERK activation. Caspase-3 was activated upon Sensor **3** treatment, which meant the apoptosis pathway was also activated in Sensor **3** treatment group. This study further confirmed that Sensor **3** specific toxicity on HER2+ cell line might be through the inhibition of proliferation pathway as well as the activation of apoptosis pathway.

DCP8a is only different from DCP8b by the position of side chain. However, the toxic properties on cell lines were completely different. DCP8a did not have toxicity while DCP8b was pretty toxic to all cell lines, which implied special interactions between

DCP8b and targets. The signaling pathways of DCP8a and DCP8b treated cells were also different. DCP8b inactivated ERK and slightly activated Caspase-3, while DCP8a only slightly inactivated pAkt. This specific toxicities might be used for better structure designs of organic probes for HER2+ breast cancer drug.

2.4 Experimental Section

2.4.1 Cell viability assay

The two MCF10A cell lines were seeded in 96-well plates (5×10^4 cells/mL, 100 μL/well). After seeding, cells were allowed to adhere and grow overnight before the addition of DCP family probes. DCP family reagents were added to the culture medium of MCF10A and cells were incubated for 21 hours. The CellTiter-Blue (G8080) from Promega (10 μL) was added to the medium (100 μL). Cells were continuously incubated for another 3 hours according to the manual. Cell viability was determined using a Spectra Max M2 Plate Reader by monitoring the absorption at 595 nm.

2.4.2 Western blots analysis

Cell lysates were prepared and equal amounts of proteins were electrophoresed on SDS-PAGE gels using standard conditions. The proteins were transferred to nitrocellulose membranes, which were probed with the following antibodies: total HER2 (Cell signaling, 4290); phosphor HER2 (Tyr 1221/1222, Cell signaling, 2243); total ERK (Cell signaling, 4695); phosphor ERK (Cell signaling, 9101); total Akt (Cell signaling, 4691); phosphorAkt (Cell Signaling, 4060); Caspase 3

(Oncogene); GAPDH (Santa Cruz, FL-335). The antibody dilutions for all detection of proteins were 1:1000.

CHAPTER 3 Surface Glycan Localization and Purification

3.1 Introduction

3.1.1 Glycosylation

Glycosylation is a critical post-translational modification (Fig. 3.1). Protein, lipids, tRNA and many secondary metabolites can be modified by glycosylation [68]. For example, glycosylation of proteins, which is important for protein folding, trafficking and stability [69], may govern biological homeostasis and affect organ development [70]. Glycoproteins comprise ~ 50% of the total cellular proteome and > 90% of the secreted proteome [71, 72]. Besides adding diversity to the protein structure, protein glycosylation may also modulate signal transduction and other functions. Glycosylation inside cells is mainly involved in intracellular signal pathways while protein glycosylation on the cell surface supports the communication and responds to extracellular stimuli [73, 74].

Alterations in the glycosylation include changes in the sites occupied by glycans and the glycan structures [75, 76]. Disruption of normal glycosylation is associated with hereditary and chronic disease such as cancer, diabetes, cardiovascular, inflammatory, neurological and neuromuscular diseases [75-77]. The glycoproteins in these diseases have been broadly studied for the purpose of diagnosis and treatment [76, 78]. However, to further investigate the function of

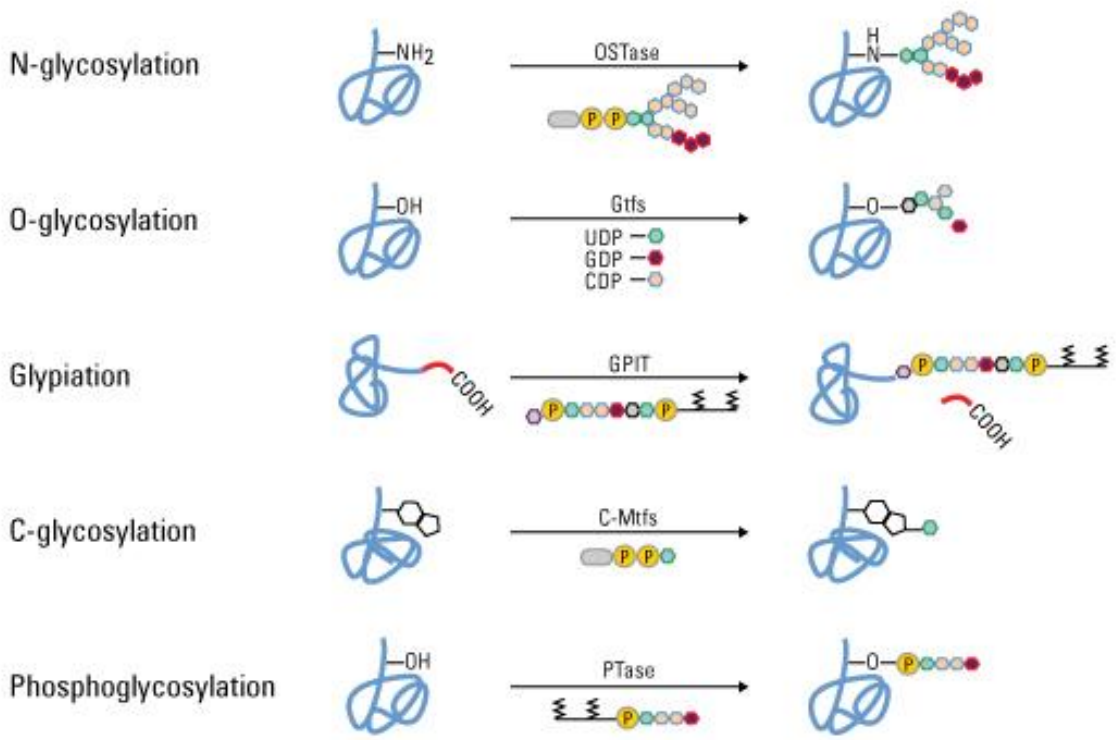


Figure 3.1. Different type of glycosylation. Figures download from the websites: http://www.piercenet.com/method/protein-glycosylation.

glycoproteins, protocols of imaging and enrichment of glycoproteins are still in demand.

3.1.2 Glycoproteomics and Glycomics

Glycoproteomics and glycomics refer to two approaches used for the characterization of specific targets' (i.e. cell, tissue or organ) glycoproteome and glycome from an extracted protein mixture (Fig. 3.2) [75, 76, 79, 80]. For glycoproteomics, glycosylated proteins were enriched with certain techniques and then analyzed by LC/MS/MS (Liquid Chromatography/Mass Spectrometry/Mass Spectrometry) for protein species and glycosylation sites [80]. For glycomics analysis, the glycans are usually released from the glycomolecules and the composition and structure are analyzed either by mass spectrometry separately or with chromatographic techniques [80].

Three different approaches were often applied to glycoproteomics: lectin affinity chromatography [81], antibody affinity chromatography [82] and metabolic labeling followed by specific interaction or reaction chromatography [83]. Depending on the methods chosen for enrichment and data analysis, the databases of glycoproteomics and glycomics might have variety among study groups. Thus, integration of all databases became more and more important. Certain programs have been established to merge databases for glycoproteomics [84, 85].

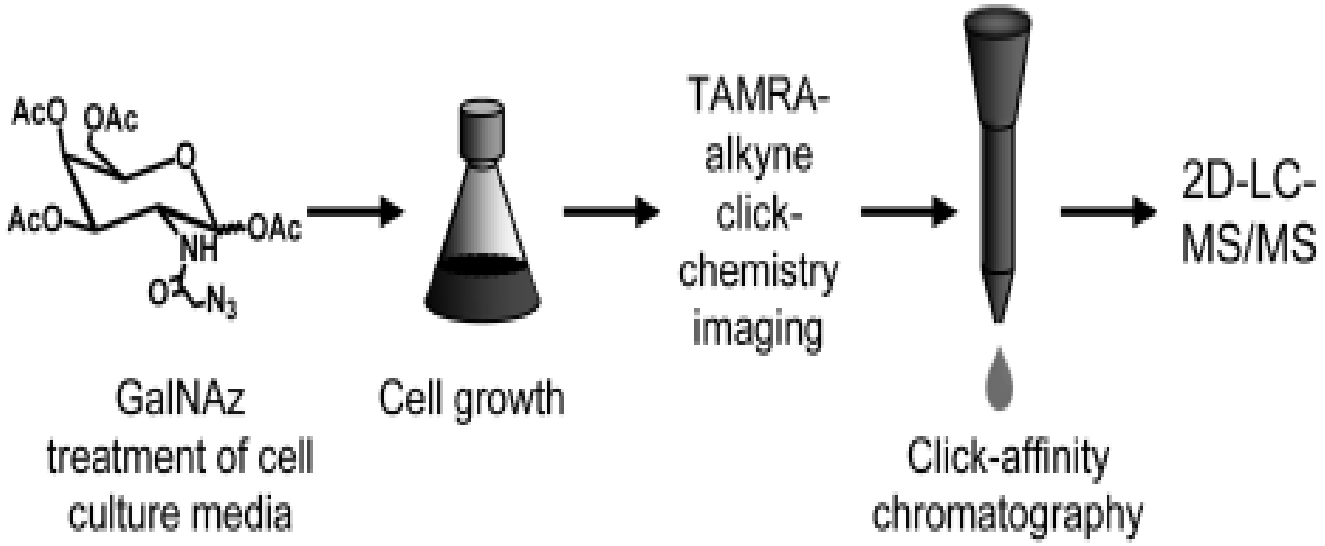


Figure 3.2. Schematic figure of protocol for glycoprotein purification using azide-alkyne cycloaddition reaction. Azide labeled non-natural sugar will be incubated with cells. After reacted with alkyne conjugated reporter, glycoprotein will be enrichment by chromatography and applied to LC/MS/MS analysis. Figure was adapted from "Slade P.G. et al, 2012, J of proteome Res." paper with copy right permission.

3.1.3 Click chemistry and its applications in bioorthogonal chemistry

Click chemistry is very often used in the imaging and enrichment of glycoproteins. Click chemistry is first reported by Dr. Sharpless group in 2001 [86]. It refers to certain reactions that is modular, wide in scope, high yields, generate inoffensive byproducts and be stereospecific [86]. One of the most famous click reactions is the azide-alkyne cycloaddition using Cu(I) as the catalyst [87].

As showed in Fig. 3.3, using this reaction and analogue reactions, some companies have developed certain kits with alkyne or azide conjugated beads to enrich glycosylated proteins [79, 88]. Fluorescence probes with azide or alkyne groups were also synthesized to determine the localization of target molecules which could be labeled by corresponding alkyne or azide [89].

Because copper might be toxic to live cells, Bertozzi and coworkers introduced two reactions that had better biocompatibility with intracellular environment and coined them as bioorthogonal reaction [90, 91]. Bioorthogonal reactions are chemical reactions that can occur inside of living systems without interfering with native biochemical processes [90]. One of the bioorthogonal reactions is called Staudinger ligation reaction, which allows azide react with triacrylphosphines (see Fig. 3.3) in live cells [92]. As first reported for bioconjugation usages in 2000 by Dr. Bertozzi's group [90], this reaction has already been applied to imaging and protein enrichment in live cells and mice [91, 93]. The other reaction is termed Strain-promoted alkyne-azide cycloaddition (SPAAC) [94]. This reaction between azide and cyclooctynes has a faster reaction rate and do not need a copper catalyst (Fig. 3.3) [88, 94].

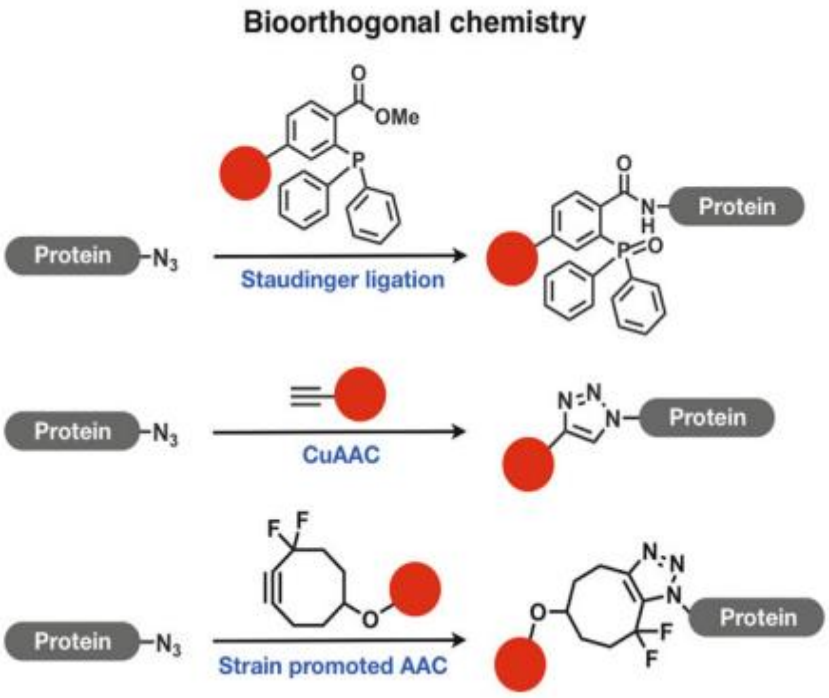


Figure 3.3. Bioorthogonal reactions for proteins labeling. Three reactions were shown here: Staudinger ligation, Cu-catalyzed Azide-Alkyne Cycloaddition and Strain Promoted Azide-Alkyne Cycloaddition reactions. Figure was adapted from book "Mass Spectrometry of Glycoproteins: Methods and Protocols, B. W. Zaro, et. al." with copy right permission.

3.1.4 Bioorthogonal chemical reporters for glycoprotein labeling

Bioorthogonal chemical reporters are non-natural molecules which could be incorporated into different biomolecules like glycan, protein and RNA *in vivo* without interrupting the molecular properties [86]. They may containing small tags i.e. azide and alkyne, which are small in size and metabolically stable [88, 95]. To study protein glycosylation, saccharides precursors tailored with such kind of tags were designed, synthesized and applied to label glycans [88]. Some typical reporters for the biorthogonal labeling of different saccharides are showed in Fig. 3.4.

Basically, those bioorthogonal saccharides precursors (Fig. 3.4) may be incorporated to glycans through different endogenous salvage pathway [96, 97]. As a result, the azide/alkyne groups can be presented on the glycosylated proteins. By reacting with the affinity reporters (e.g. biotin motif or microbeads), glycoproteins could be enriched for downstream analyses like Western-blot, in-gel imaging and mass spectrometry [73, 88, 91, 96, 98]. Labeled glycoprotein can also be imaged in live cells or animals with fluorescent reporters [92]. For example, using a DIFO-Alexa Fluor 488 conjugate, Bertozzi's group has successfully imaged the glycosylation on the cell surface [98].

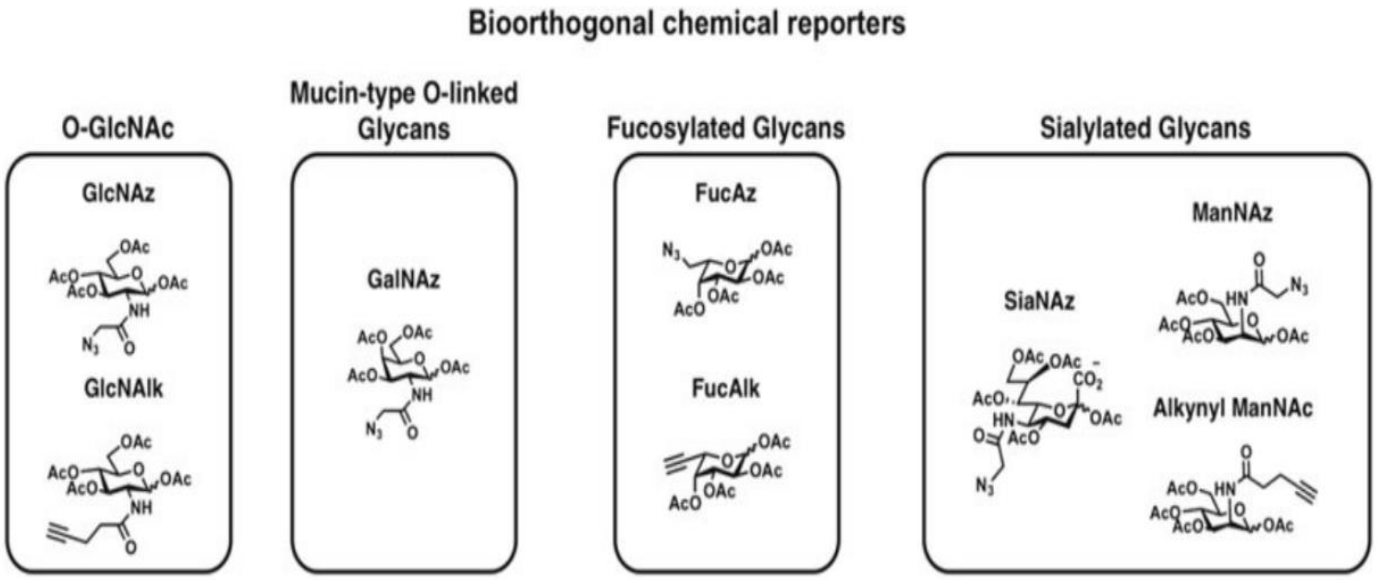


Figure 3.4. Structures of bioorthogonal chemical reporter for glycosylation. Figure was adapted from book “Mass Spectrometry of Glycoproteins: Methods and Protocols, B.W.Zaro, et. al.” with copy right permission.

3.1.5 Applications of bioorthogonal chemical reporters in glycosylation studies

Bioorthogonal reporters have variety applications in the study of glycosylation. Among them, protocols of glycoproteins imaging and enrichment are well established [79, 88, 95, 98, 99].

Different fluorescent probes have been used to image glycoproteins [88, 95, 98]. For example, quantum dots, which were great fluorescent reporters with less photo bleach and increased brightness [100], have been applied to the imaging of glycosylation. As it could be modified in variety ways, this technique had not only been applied to the imaging of glycoprotein [101, 102] but also used to interrupt glycoprotein interactions [103, 104]. Another imaging technique using dibenzocyclooctyl conjugated fluorescent probes was also broadly studied. Although different techniques were applied to the imaging of glycosylation, no detailed reports were published to compare the effect of those staining protocols. Depending on the properties of fluorescent probes, different information of glycosylation may be provided by those protocols. To image glycoproteins in cell compartments, two fluorescent probes including dibenzocyclooctyl-fluorescein and dibenzocyclooctyl-quantum dots were applied to breast cancer cell lines in this chapter. Both the total glycosylation and the membrane specific glycosylation were tracked using these two probes. The experiments provided versatile information on the localization of glycoproteins in breast cell models.

Recently, purification of bioorthogonal reporter labeled glycoprotein has become commercially available [79, 83, 99]. However, the nonspecific binding of proteins is always a challenge in those assays. To develop a better enrichment

protocol for glycoprotein, our lab synthesized Si-alkyne beads with a disulfide bonds, which could be used to release glycoproteins after enrichment. This may minimize the background from non-specific binding as well as increase the recovery of glycoproteins by this specific releasing procedure using disulfide [99].

3.2 Results and Discussion

3.2.1 Structures of fluorescent probes

Two different fluorescent probes were used in our study. One was DBCO-fluorescein by conjugating fluorescein to dibenzocyclooctyl (Fig. 3.5.a). The other is called DBCO-QT which used CdSe/CdZnS quantum dots as fluorescence cores and coated by polymers (PEG: Imidazole: Amine = 4:4:2) functionalized with dibenzocyclooctyl (Fig. 3.5.b). This DBCO-QT was designed and synthesized by our lab in collaboration with Dr. Andrew Greytak's lab (Department of Chemistry and Biochemistry, University of South Carolina). The properties of fluorescent cores determined the localization and the detection limits of the fluorescent probes. DBCO-fluorescein, as a small organic fluorescent probe, can diffuse into cells and react with targets all over the cells. Thus, it may image the total glycoproteins. As the diameter of quantum dots is around 3 nm, it can hardly diffuse into cells, which lead to the reaction sites limiting to the cell surface. Also the strong fluorescence from quantum dots helped to bring up the weak signal from cell membrane. The schematic labeling process is showed in Fig. 3.6.

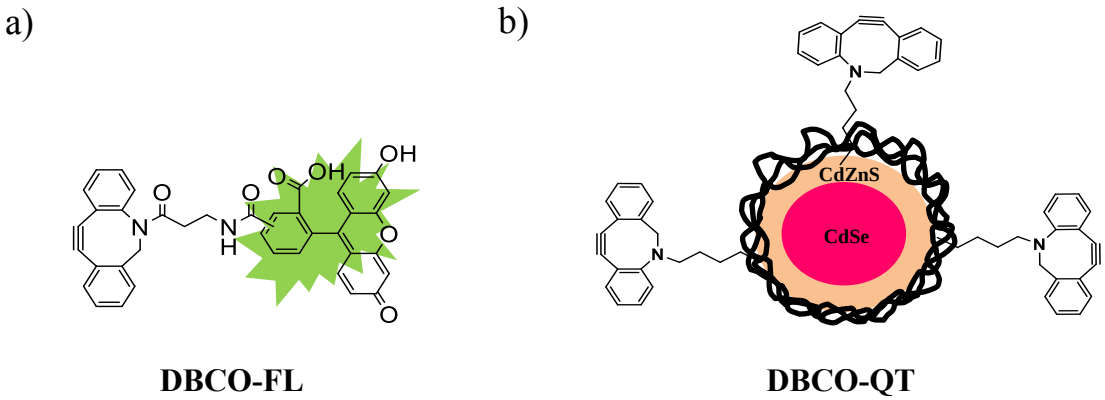


Figure 3.5. Structures of fluorescent probes for labeled glycoproteins. Structures of a) DBCO-FL (Dibenzocyclooctyl-fluorescein) and b) DBCO-QT (Dibenzocyclooctyl-quantum dots).

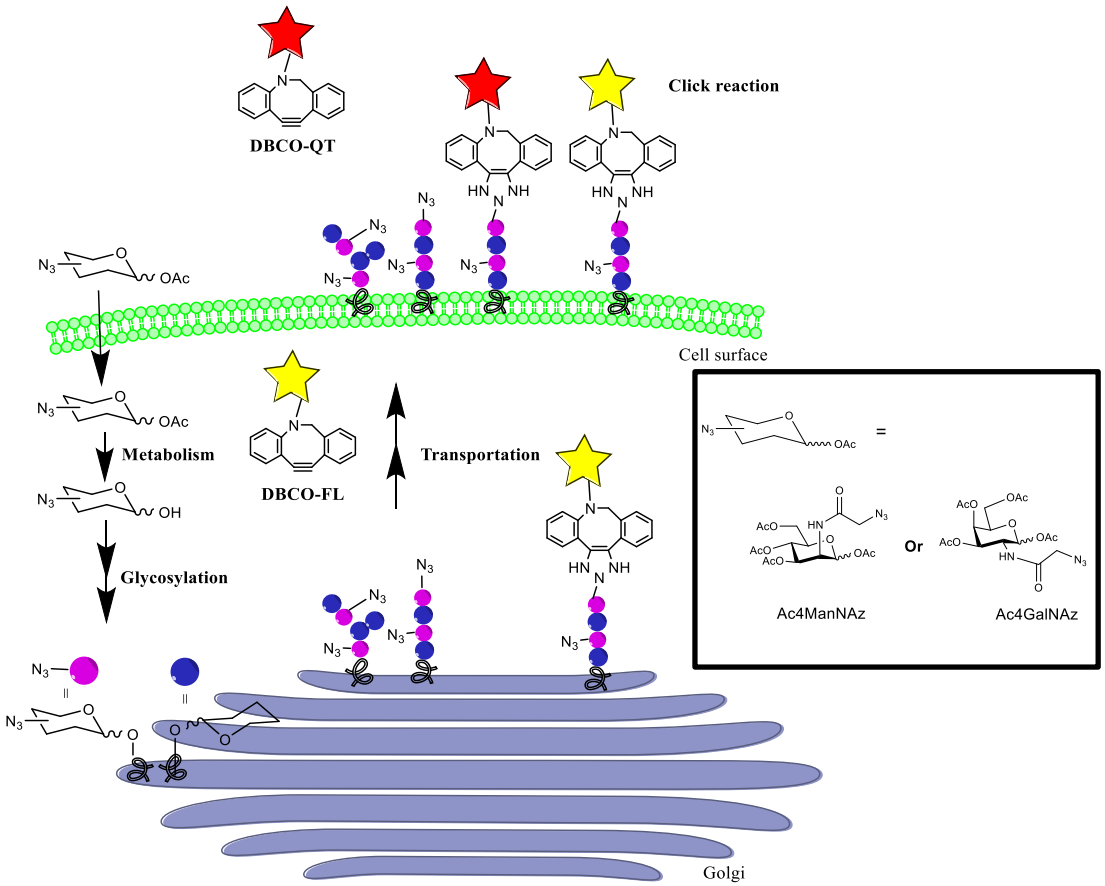


Figure 3.6. The schematic illustration of metabolic labeling and detection of glycoproteins with fluorescent probes. Cells were first fed with bioorthogonal chemical reporter Ac4ManNAz or Ac4GalNAz for certain time. Fluorescent probes were then incubated with cells for 0.5 hour in dark. Images were taken right after the staining.

3.2.2 Staining results of DBCO-fluorescein

N-azidoacetylmannosamine-tetraacylated (Ac4ManNAz) and N-azidoacetylgalactosamine-tetraacylated (Ac4GalNAz) could be processed by the N-acetyl-D-mannosamine (ManNAc) and N-acetylgalactosamine (GalNAc) salvage pathways and incorporated into glycosylated proteins [96, 105, 106]. Ac4ManNAz labeled those glycoproteins with sialylated glycans [106]. Ac4GalNAz was mostly incorporated as GalNAz for O-glycosylation [96]. Since sialylated glycans were always at the end while GalNAc was the first saccharide in O-glycosylation [107], Ac4ManNaz might be incorporated to the end of glycans and Ac4GalNAz might be the first one.

Glycosylation of proteins was different for cancer and normal cells. Not only more sialic acids were incorporated into glycans, but also sialylated glycoproteins were very often overexpressed in cancer cells [108-110]. With the Ac4ManNAz and Ac4GalNAz synthesized by Dr. Li Cai (Division of Mathematics and Science, University of South Carolina, Salkehatchie), we successfully labeled and imaged glycoproteins in MDA-MB-231 cell lines [111]. Because breast cancer cells might have higher level of sialylated proteins compared to normal cells, a higher level of cell membrane localized fluorescence was expected in cancer cells labeled with Ac4ManNAz. To prove this, another two breast cell lines, MCF10A and MCF7 were chosen. These three cell lines could be used to represent for different stages of breast cancer, normal state (MCF10A), proliferation state but not metastasis (MCF7) and metastasis state (MDA-MB-231). As cancer cell lines, MCF7 and MDA-MB-

231 should have higher level of fluorescent signals on the cell membrane for ManNAz labeled cells.

To detect the total glycosylation, the cell lines were labeled with 100 μM Ac4ManNAz for different time (1 day, 2 days, 3 days). The bioorthogonal reporter labeled glycosylation was imaged by reacting with DBCO-fluorescein (Fig. 3.7). Results showed that labeling time did affect the fluorescence intensity. The cells from Day 1 had higher fluorescence signal than those from day 2 and day 3, which might be due to the degradation of Ac4ManNAz. The fluorescent signals were spread all over the whole cells without preference on any individual organelle. To compare the staining intensity of different cell lines, three cell lines were labeled and stained parallel. Fig. 3.8 showed MCF7 and MDA-MB-231 had higher fluorescence intensity than MCF10A. This result further confirmed the previous findings on the overexpression of sialic acid in cancer cells [108-110].

3.2.3 Staining results of Quantum dots-DBCO

Quantum dots as nanometer sized particles were a new class of fluorescent probes for biomolecular and cellular imaging [101, 102]. Quantum dots have been considered as ideal imaging probes due to many unique properties including the size-tunable light emission, improved signal, simultaneous excitation of multiple colors and resistant against photo bleaching [100]. Yi Shen from Dr. Andrew Greytak's lab (Department of Chemistry and Biochemistry, Columbia, University of South Carolina) synthesized the polymer coating CdSe/CdZnS quantum dots conjugated with DBCO based on their recently published protocols.

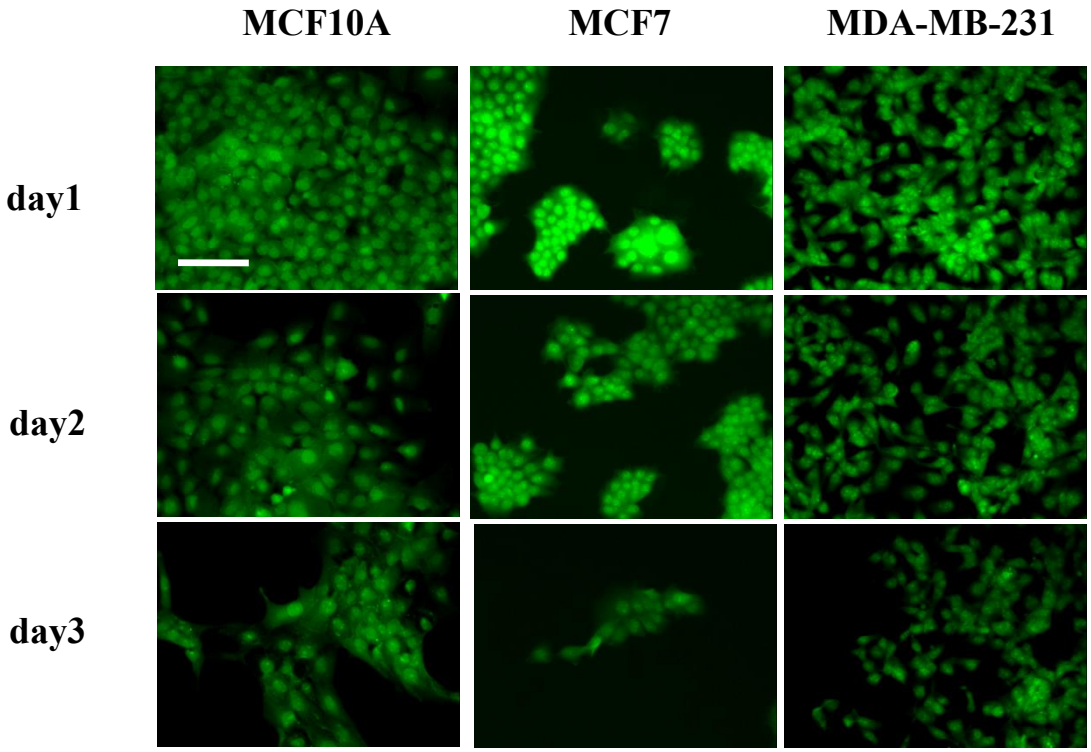


Figure 3.7. Staining of glycoprotein using DBCO-FL. MCF10A, MCF7 and MDA-MB-231 cells were incubated with 100 μM Ac4ManNAz for 1, 2 and 3 days. The cells were then fixed and reacted with DBCO-FL. Labeling time lead to different intensity of fluorescence staining. The 1 day time point was used for later experiment for the incubation time of bioorthogonal chemical reporters. Scale bar: 100 μm.

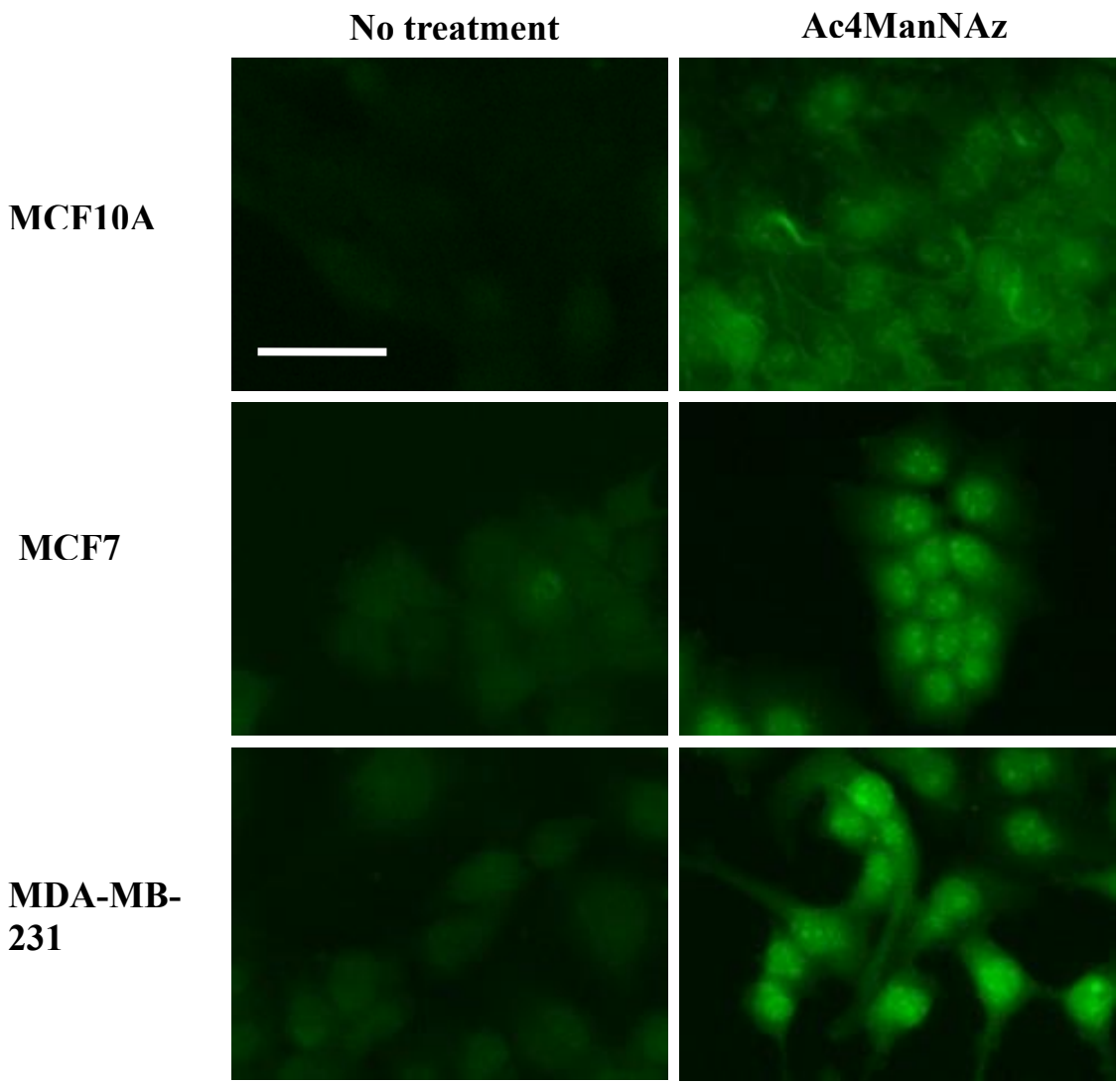


Figure 3.8. Staining of glycoprotein using DBCO-FL. MCF10A, MCF7 and MDA-MB-231 cells were incubated with 100 μM AC4ManNAz for 1 d. The cells were then fixed and reacted with DBCO-FL for 0.5 h at room temperature. The experiment was performed at the same time for all three cell lines. Identical exposure time was applied to compare the intensity of fluorescence among cell lines. Scale bar: 50 μm.

The diameter of quantum dots was around 3 nm, which made the quantum dots hard to diffuse into cells, resulting in a preferred interaction with cell surface glycoproteins. As shown in Fig. 3.9, using DBCO-QT, sialylated glycoproteins labeled by AC4ManNAz on the cell surface could be probed for three breast cell lines. While no significant fluorescent signal was found for MCF10A, both MCF7 and MDA-MB-231 shown intense fluorescence signals on the cell membrane. Therefore, the imaging method using quantum dots provided a potential way to detect and quantify the level of cell surface glycosylation.

3.2.4 Structure and stability of Si-Alkyne probe for glycoprotein enrichment

Besides the localization of glycoproteins, the compositions and structures of glycoproteome are also important for the studies of glycosylation in cancer. Current methods for glyco-molecules enrichment may have strong background from non-specific binding to beads. To minimize the non-specific binding and increase the recovery rate, our lab synthesized a Si-bead conjugated with a disulfide bond and a triple bond (Fig. 3.10). The stability of this probe could be identified by the fingerprint peaks of two carbonyl groups using infrared spectroscopy (Shimadzu FTIR Spectrophotometer). Since the click-reaction needed longer incubation, we first evaluated the stability of beads in click-reaction buffer **7** (Table 3.1). The results in Fig. 3.11 showed the beads (and their coatings) were in this buffer for as long as 3 days.

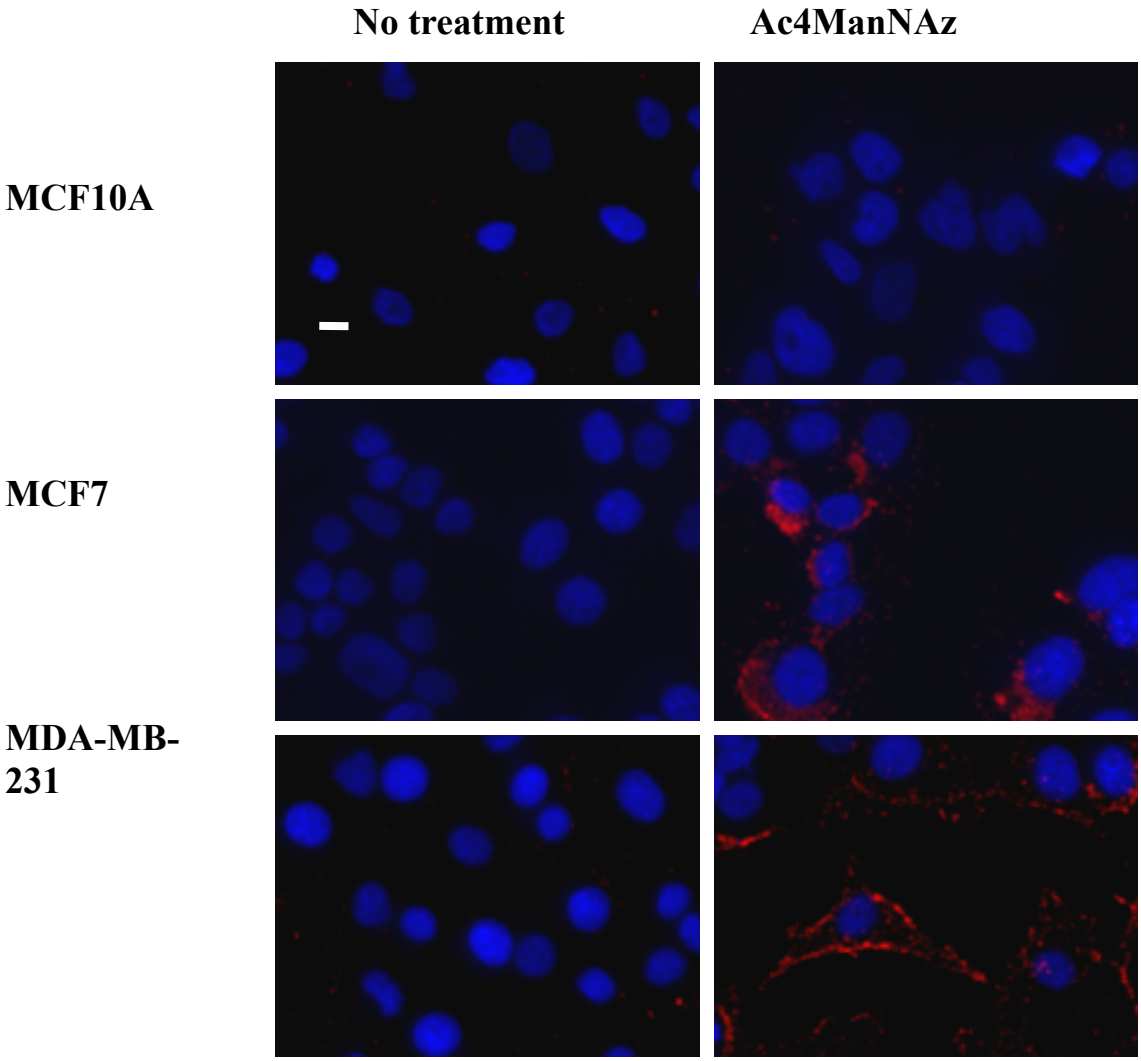


Figure 3.9. Staining of glycoprotein using DBCO-QT. MCF10A, MCF7 and MDA-MB-231 cells were incubated with 100 μM AC4ManNAz for 1 d. The cells were then fixed and reacted with DBCO-QT for 0.5 h at room temperature. The experiment was performed at the same time for all three cell lines. Identical exposure time was adopted to compare the intensity of fluorescence among cell lines. DAPI (4',6-diamidino-2-phenylindole) staining was also performed to show the localization of nucleus. Scale bar: 10 μm.

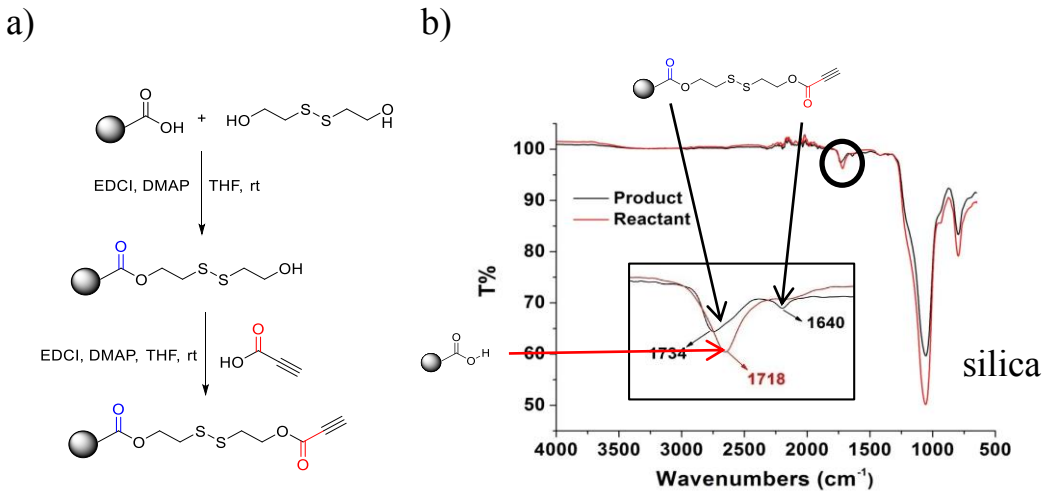


Figure 3.10. Schematic pictures to show the synthesis and characterization of Si-beads. a) The organic pathway to modify the Si-beads. b) The IR spectrum used to characterize the modification of the beads.

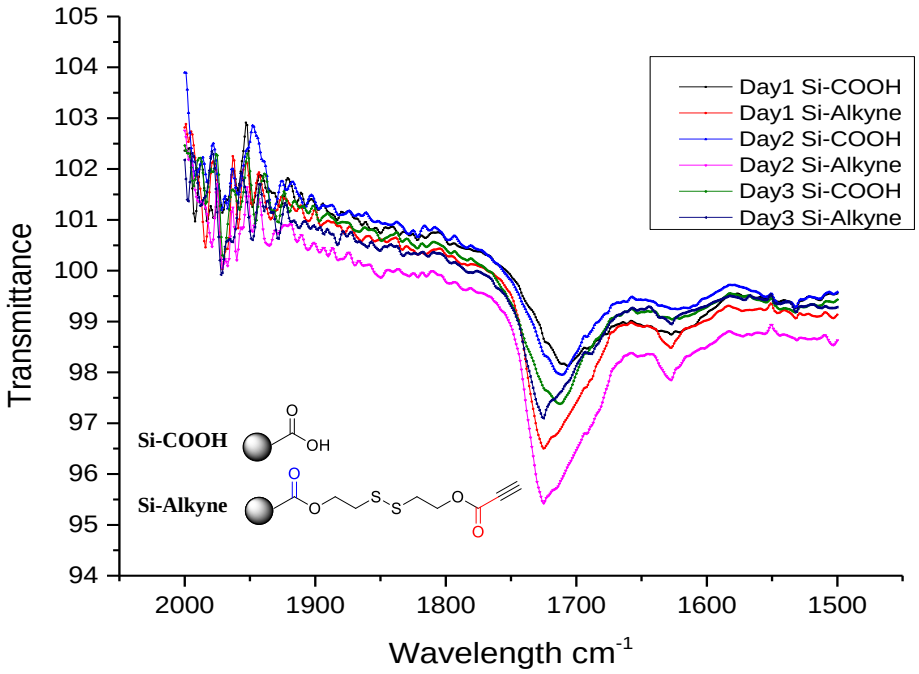


Figure 3.11. Stability of S-S bonds in buffer **7**. Beads were incubated with buffer **7** for 1 to 3 d. After washed with 100 mM Tris-HCl pH 8 buffers, beads were dried and the IR was read. The disulfide bond could be stable in buffer for up to 3 d.

Table 3.1. Buffers used in enrichment of glycoprotein with Si-beads

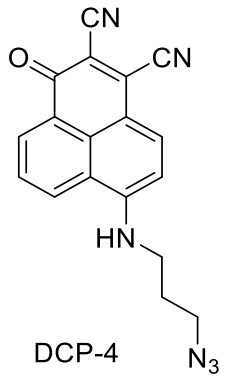
Buffer	**Property**	**Composition**
Buffer **1**	Urea lysis buffer	8 M urea, 200 mM Tris-HCl pH 8.0, 4% CHAPS, 1 M NaCl, 2 × protease inhibitor
Buffer **2**	NP40 lysis buffer	50 mM Tris-HCl pH 7.4, 150 mM NaCl, 5 mM EDTA, 1% NP-40
Buffer **4**	Wash buffer	2% SDS, 1% NP-40, 150 mM NaCl, 1% NaDeoxycholate, 50 mM Tris-HCl pH 7.5
Buffer **5**	Click reaction stock	2 mM L-Histidine, 2 mM CuSO4, 20 mM Sodium Ascobic Acid, 83.5 mM Tris-HCl pH 8.0
Buffer **7**	Click reaction buffer	1:1 mixture of buffer **4** and buffer **5**.
Buffer **9**	Wash buffer	100 mM Tris-HCl pH 8.0, 1% SDS, 250 mM NaCl, 5 mM EDTA
Buffer **10**	Wash buffer	8 M urea and 100 mM Tris-HCl pH 8.0
Buffer **11**	Wash buffer	20% acetonitrile

3.2.5 Efficiency of glycoprotein enrichment

The enrichment of glycoproteins was involved in the conjugation reaction between bioorthogonal labeled protein and specific modified beads. The cell lysates with Ac4GalNAz labeled glycoproteins were incubated with the beads in click reaction buffer. Azide tag on the glycoproteins was reacted with the alkyne on the beads and formed a 1,2,3-triazole ring. After conjugated to the beads, the glycoproteins could undergo stringent wash steps and finally be released from the beads either by reducing agents to break the disulfide bond or trypsin digestion. To control for the efficiency of each steps for glycoproteins enrichment, a fluorescent dye DCP4, which mimic proteins labeled by Ac4GalNAz, was incubated with the beads. DCP4 had an azide group that could be reacted with the triple bonds through cupper (I) catalyzed-azide alkyne cycloaddition reaction. Enrichment products could be tracked by the fluorescence of DCP4 using plate reading.

Samples from each step were collected and the fluorescent intensity was read on SpectraMax M2 with excitation at 560 nm and emission at 600 nm. As we expected, fluorescent signals in washing buffers decreased from 18698 to 2000 after 20 times washing (Fig. 3.12) for Si-Alkyne beads. It was similar for Si-COOH control beads, of which the fluorescent signal dropped faster (from 12854 to 39.8 after 7 washes). This decrease in fluorescence was due to a reduction in the non-specific binding of dyes to the beads. Fluorescent intensity was comparable for the three times of elution with 50 mM DTT for Si-Alkyne (6150, 4374, 4514). This low efficiency in elution might be due to the hydrophobicity of DCP4. After reacting with the triple bond, DCP4 might wrap around the disulfide bond and kept it from interacting with DTT.

a)



b)

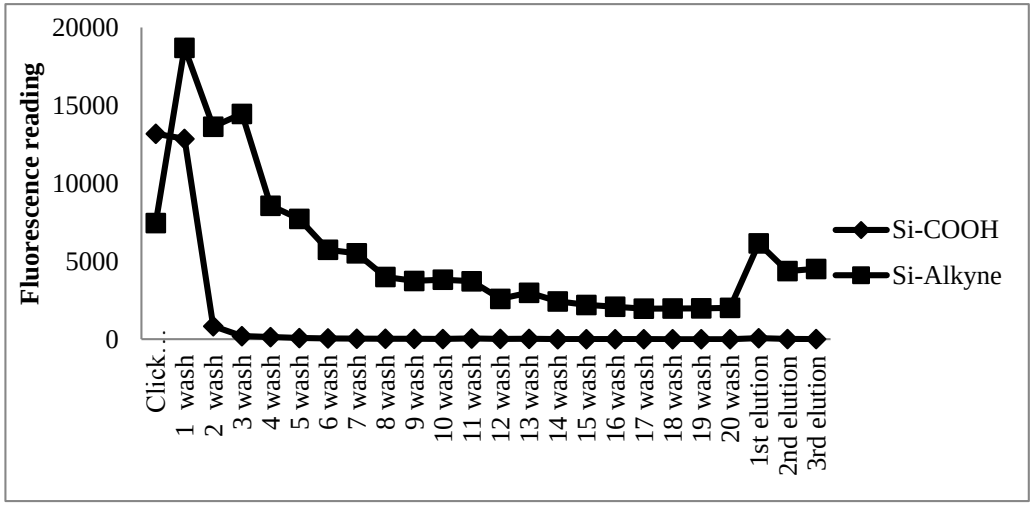


Figure 3.12. Purification control for the enrichment of glycoproteins. a) Structure of the DCP4 fluorescent probe. b) The fluorescent intensity of solution at each steps of process. Stringent wash could help to minimize non-specific binding of dyes. Elution with protein loading buffer and 50 mM DTT could release DCP4 dyes from the beads.

3.2.6 Glycoproteins enrichment with Si-Alkyne beads

To optimize the enrichment protocols, two different lysis buffers and two different releasing methods for glycoproteins enrichment were tested by the same protocol mentioned above. Cells were labeled for 1 day using 100 μM GalNAz prior to collection. Both Urea Buffer (Fig. 3.13) and NP-40 buffer (Fig. 3.14) could lyse 90% of cells. After enrichment and recovery of glycoproteins, no specific protein bands were visualized from DTT elution. To double check the elution efficiency, trypsin digestion was applied to collect proteins on the beads. However, trypsin did not help to improve the results (Fig. 3.15). One explanation was that the labeled proteins might not react with the beads efficiently. We might need to modify and improve the properties of the beads to allow higher efficiency in click reaction between beads and proteins.

3.3 Conclusion

To compare the glycosylation level among three breast cell lines, the biorthogonal chemical reporters, i.e. AC4ManNAz and AC4GalNAz, were used to labeled cells and two different detection probes were applied to track glycoproteins. DBCO-fluorescein could diffuse into cells while DBCO-QT predominantly recognizes azide tags on the cell membrane. Total glycosylation level among cells could be analyzed by DBCO-fluorescein while the cell membrane localized glycoproteins could be detected by DBCO-quantum dots. Both total glycosylation and cell surface glycosylation level were higher in the two breast cancer cell lines MCF7 and MDA-MB-231 compared to normal control cell lines.

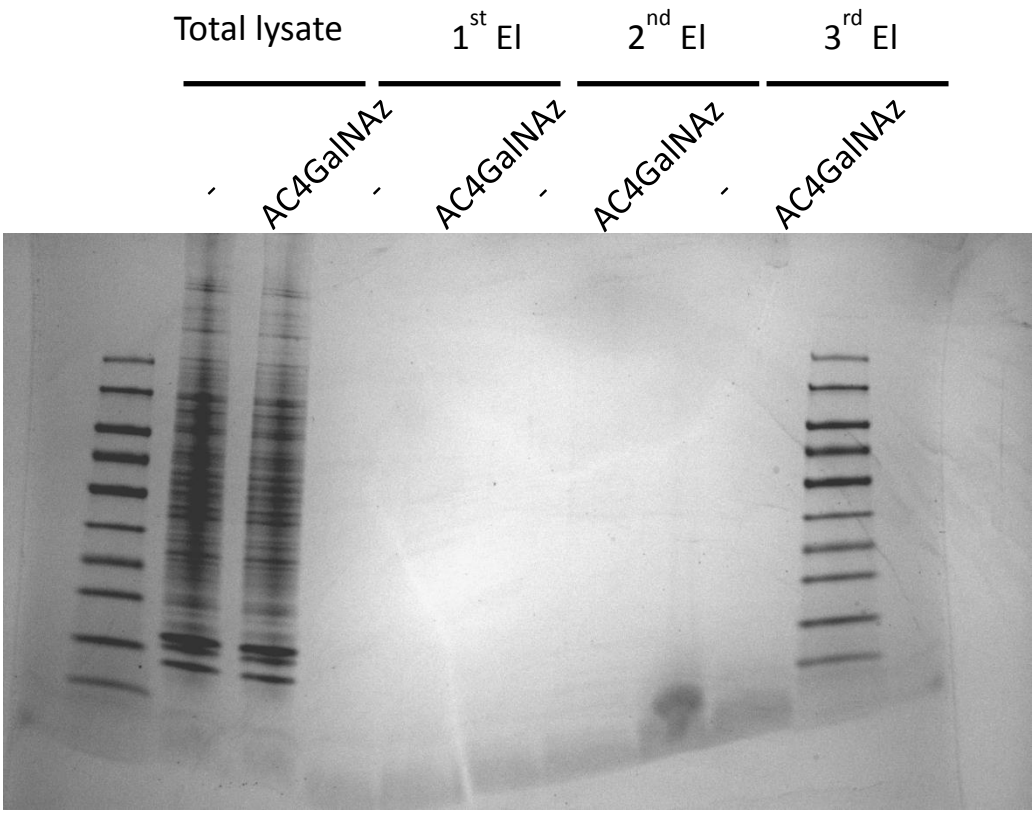


Figure 3.13. Enrichment of glycosylated protein using Si-Alkyne beads: choice of lysis buffer. Urea lysis buffer was used to lysis HEK293T cells labeled with 100 μM GalNAz for 1 d. Proteins were harvest and incubated with beads for 1 d in click reaction buffer. Proteins were eluted in loading buffer plus 200 mM DTT. No significant enrichment protein bands were seen for labeled cells compared to unlabeled cells.

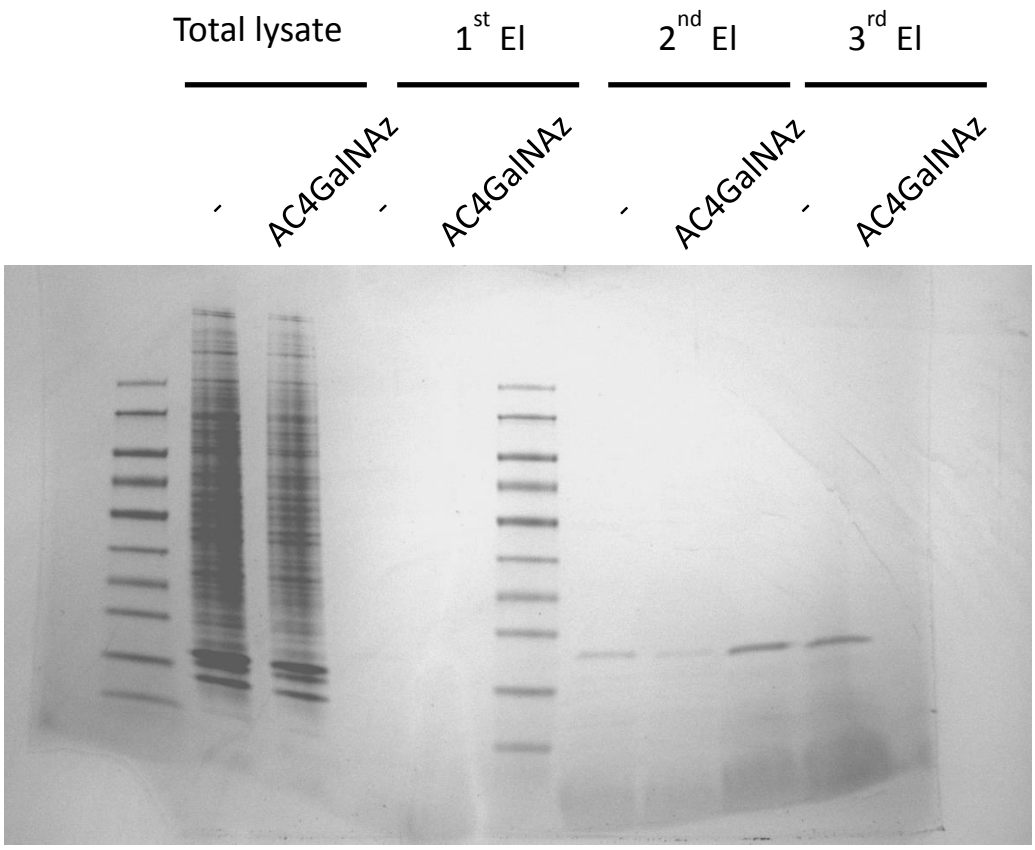


Figure 3.14. Enrichment of glycosylated protein using Si beads-choice of lysis buffer. NP-40 lysis buffer was used to lysis HEK293T cells labeled with 100 μM GalNAz for 1 d. Proteins were harvest and incubated with beads for 1 d in click reaction buffer. Proteins were eluted in loading buffer plus 200 mM DTT. No significant enrichment protein bands were seen for labeled cells compared to unlabeled cells.

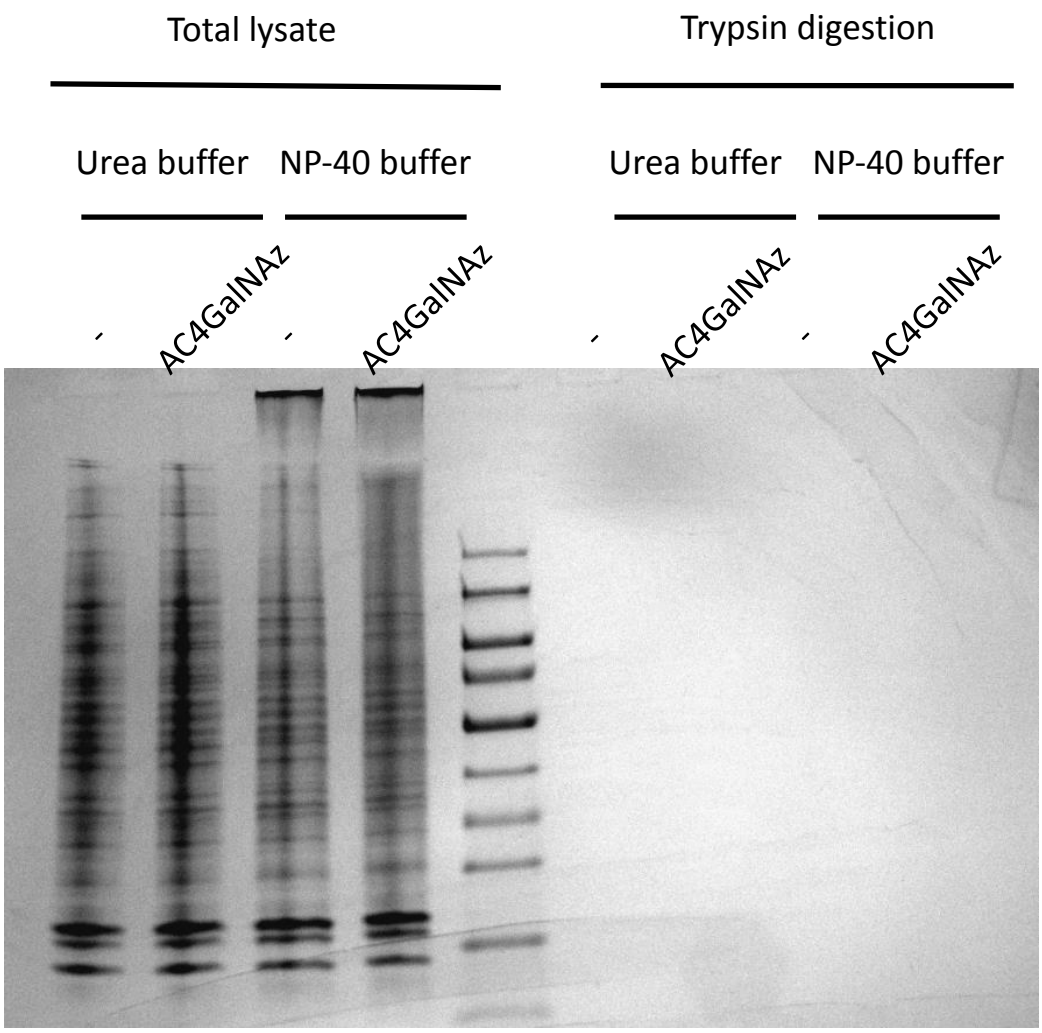


Figure 3.15. Enrichment of glycoprotein using Si beads-choice of protein releasing methods. Urea buffer and NP-40 lysis buffer were used to lysis HEK293T cells labeled with 100 μM GalNAz for 1 d. Proteins were harvest and incubate with beads for 1 d in click reaction buffer. Proteins were digested on beads by trypsin. No significant enrichment protein bands were seen for labeled cells compared to unlabeled cells.

Since Si is an inorganic reagent, the nonspecific bindings between biomolecules and the beads could be reduced. This may help to improve the efficiency of protein enrichment. To achieve this research purpose, our lab synthesized the si-probes with disulfide bond and triple bond. The disulfide bond could be used to release the proteins from the beads after enrichment procedures. This design could obtain the intact proteins as well as intact glycans instead of peptides fragments (from trypsin treatment). The stability of this bead was tested in a combination of washing buffer and click reaction buffer. The disulfide bonds were still stable after 3 days incubation. To further confirm the efficiency of each purification step, a control experiment using a fluorescent probe to mimic labeled glycosylated proteins was performed. The fluorescent signals from control experiment confirmed the click reaction was successfully reacted. The DCP dyes conjugated to the beads could then be release by DTT treatment to break the disulfide bond. Although we confirmed the recovery of the controlled fluorescence dyes, we did not succeed in developing enrichment protocols for glycoprotein. This might be attributed to the low accessibility of proteins to the triple bounds, i.e. the linker was not long enough to give high reaction efficiency between triple bond and azide group on labeled proteins. We could try beads with longer linker later for this project. We might also coat the beads with polymers to help the protrusion of the triple bond end.

3.4 Experimental Section

3.4.1 Cell Maintenance

The MCF10A and MDA-MB-231 cell lines were kindly gifts from Prof. Hexin Chen (University of South Carolina, Department of Biology Science). MCF7 cell line was

purchased from American Type Culture Collection (Cat. No: HTB-22). MCF10A were maintained in DMEM:F12 (1:1) medium with 5% horse serum, 10 μg/mL insulin, 20 ng/mL EGF, 100 ng/mL cholera toxin, 0.5 μg/mL hydrocortisone. MCF7 cells were maintained in DMEM-high glucose medium with 10 μg/mL Insulin and 10% FBS. MDA-MB-231 was cultured in DMEM-high glucose medium with 10% FBS. Please see Chapter 1 for HEK293T cell line information. All cells were in a water-saturated atmosphere of 5% CO2 and 95% air at 37 ℃.

3.4.2 Staining of cells with DBCO-FL and DBCO-QT

Cells were seeded at 1×10^5 cells/well in the 12-well plates and allowed to grow overnight. The Ac4ManNAz and Ac4GalNAz were added to the cells the next day and incubated for indicated time. On the day of imaging, cells were washed three times with PBS and fixed by 4% paraformaldehyde for 15 min at room temperature. After being washed three times with PBS, the samples were blocked in 100 mM glysine for 30 minutes. DBCO-FL (10 μM final concentration) or DBCO-QT (20 nM final concentration) were incubated with cells for 30 minutes. Cells were washed three times and kept in PBS in dark before imaging. DAPI were added to the PBS at 5 μg/mL final concentration without washing. Confocal images were taken by Olympus X81 fluorescence microscopy.

3.4.3 Stability of Si-beads

Si beads were incubated in the buffer **7** for indicated time. After washed with PBS for three times, the beads were air-dry overnight. The infrared spectra were

recorded with a Shimadzu FTIR Spectrophotometer. Buffer compositions were showed in Table 3.1.

3.4.4 Beads preparation and reactivity with DCP4

One day before protein collection, beads were blocked with 0.0001% poly vinyl alcohol in 100 mM Tris-HCl pH 8.0 overnight at 4 ºC. After washed three times with 100 mM Tris-HCl pH 8.0, 15 mM (around 15 mg) beads was incubated with 5 mM DCP-4 for 18 hours in Buffer **7**. The beads were then pellet and washed five times with Buffer **9**, ten times with Buffer **10** and five times with Buffer **11**. Then the dyes that conjugated to the beads were eluted by boiling the beads with protein loading buffer (BioRad, Cat. No: 161-0737) plus 50 mM DTT for 5 min. The washing buffer and elution solution were collected and pipetted into 96 well to read the fluorescence (excitation at 560 nm and emission at 600 nm).

3.4.5 Protein purification and SDS-PAGE verification

HEK293T cells were seeded at 3×10^6/plate in 10 cm plate. Each sample needs 8 plates in order to collect 5×10^7 cells. Cells were dissociated by 0.5 mM EDTA in PBS for 10 minutes. Harvested cells were subjected to lysis buffer (Urea buffer or NP40 buffer) and homogenized using Pro Scientific hand-held homogenizer (PRO-01-02200 PRO200). The total cell lysates from NP40 buffer were further cleared by methanol chloroform precipitation and resolved in lysis buffer II. Protein concentrations were measured using Thermo Commasie Plus Protein reagents. Equal amount of proteins were subject to beads pre-blocked with PVA (polyvinylalcohol) and incubated with Buffer **5** for 18 hours. The beads were then pellet and washed five times with Buffer **9**, ten times

with Buffer **10** and five times with Buffer **11**. Then the glycosylated proteins that conjugated to the beads were eluted by boiling the beads with protein loading buffer 200 mM DTT for 5 minutes and briefly centrifuged. For trypsin digestion, 0.005 μg/μL trypsin (final concentration) was incubated with beads in the digestion buffer (100 mM Tris, 2 mM $CaCl_2$, 10% acetonitrile) for 8 hours at 37 ℃. The digestion solutions were collected by briefly pellet the beads. Cell lysates and eluting solutions were prepared. Equal amounts of protein were electrophoresed on SDS-PAGE (4%-10%). The gels were stained with Lagsafe Gel Blue (Cat. No 786-35) and images were taken by Bio-Rad Gel imaging system (GelDoc XR).

CHAPTER 4 Assessment of Aldehyde Dehydrogenase 2 (ALDH2) in Live Cells

4.1 Introduction

4.1.1 ALDH family proteins

Aldehydes are highly reactive electrophilic compounds that play a vital role in physiological process such as mutagenic, carcinogenic and cytotoxic detriment [112, 113]. Various reactive aldehydes can be produced from different endogenous precursors [114]. Aldehydes can be further oxidized to acetic acid and detoxified by aldehyde dehydrogenase (ALDH) [112] (Table 4.1). The *ALDH* family currently has 19 members (Table 4.1) [115]. Due to their important functions in aldehyde metabolism, mutations and polymorphisms of ALDH genes (loss of function) are often found to be associated with different diseases in humans and rodents [113]. In particular, polymorphisms of ALDH2 (a subtype of ALDH) that resulted in loss of function of ALDH2, were associated with an increased risk for various cancers, including oropharyngolaryngeal, esophageal, stomach, colon, lung, head and neck cancers [116, 117].

4.1.2 Cancer stem cell model

Cancers are heterogeneous, multicellular entities that arise when an epigenetic or genetic change occurs in normal cells, which disrupt normal cellular homeostasis [118, 119]. Currently there are two models of heterogeneity for solid cancer cells [120] (Fig. 4.1). One model believes cancer cells include many different subtypes

Table 4.1. Human ALDH family proteins [121]

ALDH isoenzyme	Cellular Localization	Tissue/Organ Distribution	Disease Associated with Polymorphisms
ALDH1A1	Cytosol	Liver, kidney, red blood cells (RBCs), skeletal muscle, lung, breast, lens, stomach mucosa, brain, pancreas, testis, prostate, ovary	Perhaps alcoholism
ALDH1A2	Cytosol	Testis, small amounts in liver, kidney	Spina bifida
ALDH1A3	Cytosol	Kidney, skeletal muscle, lung, breast, testis, stomach mucosa, salivary glands	
ALDH1B1	Mitochondria	Liver, kidney, heart, skeletal muscle, brain, prostate, lung, testis, placenta	Various phenotypes
ALDH1L1	Cytosol	Liver, skeletal muscle, kidney	
ALDH1L2	Unknown		
ALDH2	Mitochondria	Liver, kidney, heart, skeletal muscle, lens, brain, pancreas, prostate, spleen	Ethanol-induced cancers, hypertension
ALDH3A1	Cytosol, nucleus	Stomach mucosa, cornea, breast, lung, lens, esophagus, salivary glands, skin	Various phenotypes
ALDH3A2	Microsomes, peroxisomes	Liver, kidney, heart, skeletal muscle, lung, brain, pancreas, placenta, most tissues	Sjögren-Larsson syndrome
ALDH3B1	Cytosol	Kidney, lung, pancreas, placenta	
ALDH3B2	Unknown	Parotid gland	
ALDH4A1	Mitochondria	Liver, kidney, heart, skeletal muscle, brain, placenta, lung, pancreas, spleen	Type II hyperprolinemia
ALDH5A1	Mitochondria	Liver, kidney, heart, skeletal muscle, brain	Γ - hydroxybutyric aciduria
ALDH6A1	Mitochondria	Liver, kidney, heart, skeletal muscle	Developmental delay
ALDH7A1	Cytosol, nucleus, mitochondira	Fetal liver, kidney, heart, lung, brain, ovary, eye, cochlea, spleen, adult, spinal cord	Pyridoxine-dependent epilepsy
ALDH8A1	Cytosol	Liver, kidney, brain, breast, testis	
ALDH9A1	Cytosol	Liver, kidney, heart, skeletal muscle, brain, pancreas, adrenal gland, spinal cord	Various phenotypes
ALDH16A1	Unknown	Neuronal cells	
ALDH18A1	Mitochondria	Kidney, heart, skeletal muscle, pancreas, testis, prostate, spleen, ovary, thymus	

which have the potential to proliferate extensively, but none can form spheres or colonies in clonogenic assays [120]. Cancer stem cell model believes that a subset of cancer cells called cancer stem cells (CSCs) can consistently proliferate extensively in clonogenic assays and form new tumors on metastasis sites [120, 122, 123].

Cancer stem cells were first isolated by Dr. John Dick in leukemia in 1994 using $CD34^{+}CD38^{-}$ as the cancer stem cell marker [124]. The first report of cancer stem cells in solid tumors came in 2003 using a $CD44^{+}CD24^{-}$ marker for breast cancer [125]. CSCs have also been identified in several other cancer types, including solid tumors of the liver, pancreas, brain, colon and prostate [118, 124-134].

4.1.3 ALDH proteins may be used as cancer stem cell marker

In 1990s, Jones and Storms [135, 136] developed a technique known today as the Aldefluor Assay to stain and isolate live hematopoietic stem cells with high ALDH activity (Fig. 4.2). Cells with high ALDH activity uptake the fluorescent substrate BODIPY-aminoacetaldehyde (BAAA) by passive diffusion and then convert BAAA into negatively-charged BODIPY-aminoacetate (BAA-). BAA- then retains inside cells, which makes the subset of cells with a high ALDH activity (ALDH-hi) become highly fluorescent. These ALDH-hi populations can be distinguished through comparison to the diethylaminobenzaldehyde (DEAB) negative control (a specific inhibitor of ALDH [137]) (Fig. 4.2.b). To date this method has been applied to different cell types to identify stem cells or cancer stem cells [138-142].

As the BAAA is distributed all over the cell, this assay basically measures the activity of all ALDHs expressed in cells. It is non-specific and may be affected by cell

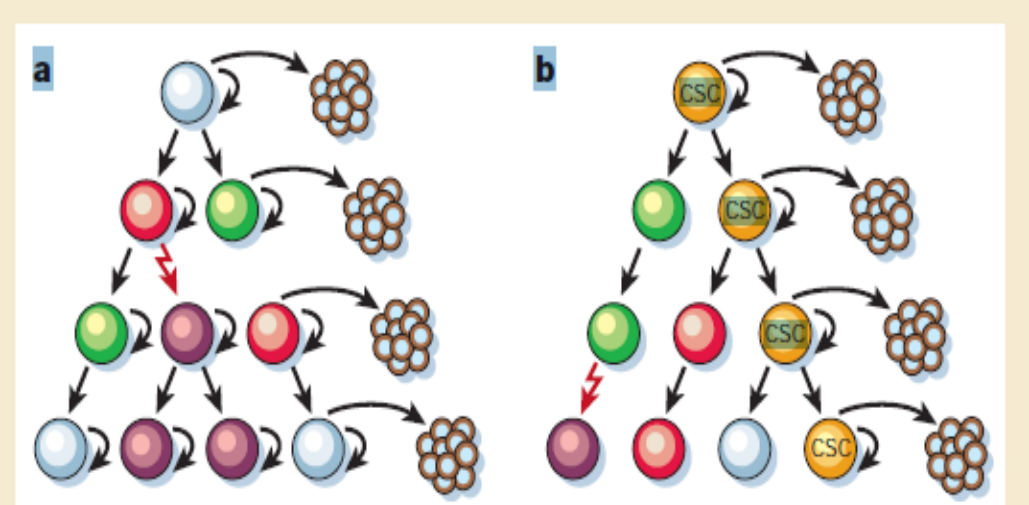


Figure 4.1. Cancer stem cell models. a) Cancer cells of many different phenotypes have the potential to proliferate extensively, but any one cell would have a low probability of exhibiting this potential in an assay of tumorigenicity. b) Most cancer cells have only limited proliferative potential, but a subset of cancer cells (cancer stem cell) consistently proliferate extensively in clonogenic assays and can form new tumors on transplantation. Figure was adapted from "Reya, T., et al., Nature, 2001" paper with copy right permission.

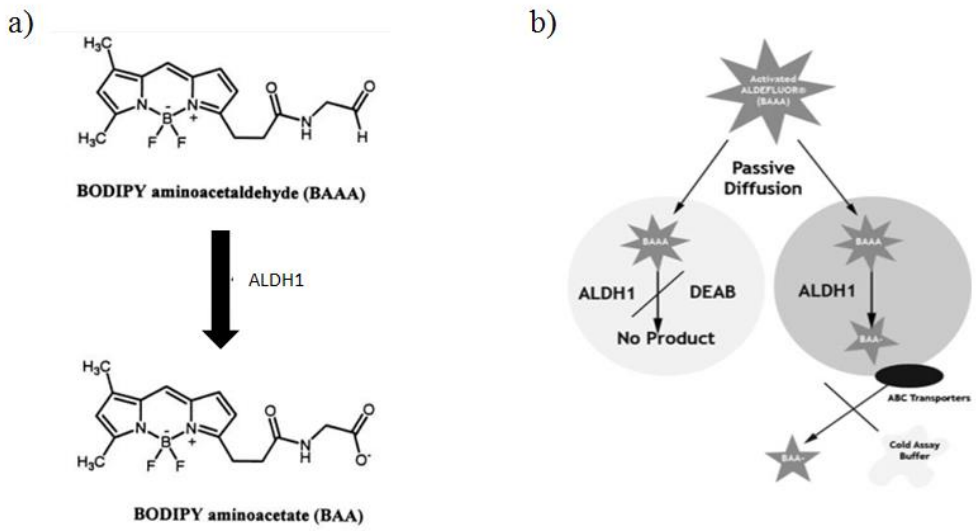


Figure 4.2. Aldefluor Assay. a) ALDH converts BODIPY-aminoacetaldehyde (BAAA) to BODIPY-aminoacetate (BAA-). b) Schematic mechanism of Aldefluor Assay. Cells expressing ALDH uptake BAAA and convert it to BAA-. BAA- is then retained inside cells, causing the subset of ALDH-Hi cells to become highly fluorescent. Cold assay buffer is added to prevent the ATP-binding cassette (ABC) transporters to pump the BAA- out of the cells. As a negative control, diethylaminobenzaldehyde (DEAB), which is a specific ALDH inhibitor, is used to quench the activity of ALDH-Hi cells and prevent the cells to become fluorescent. Figure was adapted from "Storms, R.W., et al, Proc Natl Acad Sci, 1999" and "Ma, I., et al., Stem Cell Rev and Rep, 2011" with copy right permission.

types. It will be interesting to know the correlations between specific members of ALDHs and cancer stemness.

4.1.4 ALDH2 proteins might be a marker for cancer stem cells

ALDH2 encodes a 50 kDa mitochondrial matrix protein that is constitutively expressed in a variety of tissues including liver, kidney, heart, lung and brain [143]. ALDH2 is the primary enzyme involved in the oxidation of acetaldehyde (Km < 0.1 μM) during ethanol metabolism [144]. Polymorphisms of ALDH2 that leads to loss of function of ALDH2 was associated with an increased risk for various cancers, including oropharyngolaryngeal, esophageal, stomach, colon, lung, head and neck cancers [116, 117]. Alcoholic individuals carrying ALDH2 (ALDH2*2, E487K) mutations displayed increased levels of acetaldehyde-derived DNA adducts, which indicated a potential role of ALDH2 mutation in DNA damage and cancer development [144]. ALDH2*2 was also associated with a lack of nitroglycerin efficacy in Chinese patients [145], the myocardial infarction in Korean patients and the hypertension in Japanese patients [146, 147]. In addition, ALDH2 played a major role in the scavenging of the neurotoxic aldehyde metabolite of dopamine, DOPAL (Km= 4.2 μM) [113, 148], which might explain why ALDH2*2 was associated with an elevated risk for late on-set Alzheimer in Japanese [149] and Chinese [150].

Because of its important enzyme activity, ALDH2 might play an important role in cancer stem cell signaling and might be identified as a marker for cancer stem cells.

4.1.5 Methods to study ALDH2 protein

Some traditional methods i.e. *in vitro* activity assay, immunohistochemistry and Western-blot have been used to study ALDH2's activity and expression in cancer [113, 117, 134, 144, 148, 151]. Unfortunately, those methods have some disadvantages. First, these methods require the endogenous ALDH2 to be released from the cells, which kills the cells [152, 153]. For cancer stem cells, identifying and isolating the live cells are required to study their function and phenotype, especially their resistance to anti-cancer drug [120]. It will be more useful to study the ALDH2 enzyme activity in viable cells. Second, cross reactivity of antibodies against different ALDHs introduced non-specific binding [154, 155]. New methods needed to be developed to identify and isolate ALDH2-high cells efficiently for further research purposes.

To investigate the possibility of ALDH2 as a cancer stem cell marker, methods detecting ALDH2 activity in live cell were needed to be developed and optimized. For this purpose, our lab synthesized two different fluorescent probes: RSAAA and RSACOOH (Fig. 4.3). Using these two probes, we developed and optimized a new method to isolate cells with higher ALDH2 enzyme activity.

4.2 Results and Discussion

4.2.1 Structures and spectra of RSAAA and RSACOOH

The structures of two fluorescence probes were present in Fig. 4.3. The schematic conversion route of these molecules was also showed. RSAAA can be *in situ* converted to RSACHO inside cells by a quick dehydration reaction. RSACHO could then work as a substrate for ALDH family proteins and be oxidized to RSACOOH in the presence of

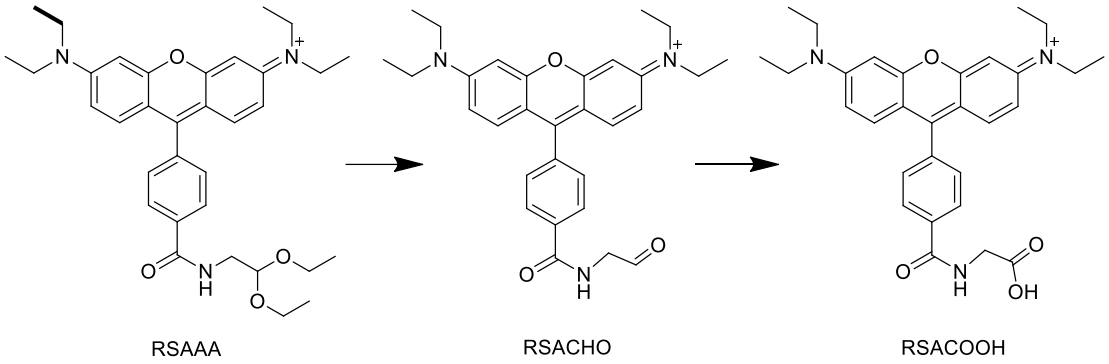


Figure 4.3. Structures of RSAAA and RSACOOH probes and their possible conversion pathways in cells. The acetal structure of RSAAA can be easily transformed into RSACHO *in situ*. RSACHO then worked as substrate for ALDH2 and was oxidized to RSACOOH.

ALDH family proteins. As RSAAA had a positive charge, we postulated that it primarily localized in the mitochondria. Conversely, the negative charge on RSACOOH would make it easy to be pumped out of the cell. Thus, the cells with higher ALDH2 activity should have a lower fluorescence.

ALDH2 is the predominant member of ALDH proteins in mitochondria [143, 144] and the major member to oxidize RSACHO. Therefore we hypothesized that the higher the activity of ALDH2, the lower the fluorescence signals (Fig. 4.4). As RSACHO was more reactive than RSAAA and might induce high background, RSAAA might be a better candidate to assay for ALDH2 activity.

To test the fluorescence activity of RSAAA and the end product RSACOOH, their spectrum was measured on Varian Cary Eclipse Fluorescence Spectrophotometer (Fig. 4.5). RSAAA had much higher fluorescence than the RSACOOH (Fig. 4.5.a). Their emission both peaked around 592 nm and their fluorescence could be tracked at the same excitation and emission range. To confirm the retention of RSAAA and RSACOOH in the cells, they were used to stained cells for 0.5 hour before the probes were removed and the cells were allowed to chase for different time. Fluorescence of cells was recorded in 96 well plates with a SpectraMax M2 Microplate Reader (excitation: 536 nm; emission: 592 nm). As shown in Fig.4.5.b, the fluorescence of RSACOOH staining cells dropped by 95% in the first 20 min after removing the fluorescence probe, while RSAAA staining cells retained 40% of fluorescence activity and eventually dropped to 20% after chasing for 115 min. These results confirmed the hypothesis on the intracellular retention of these two dyes.

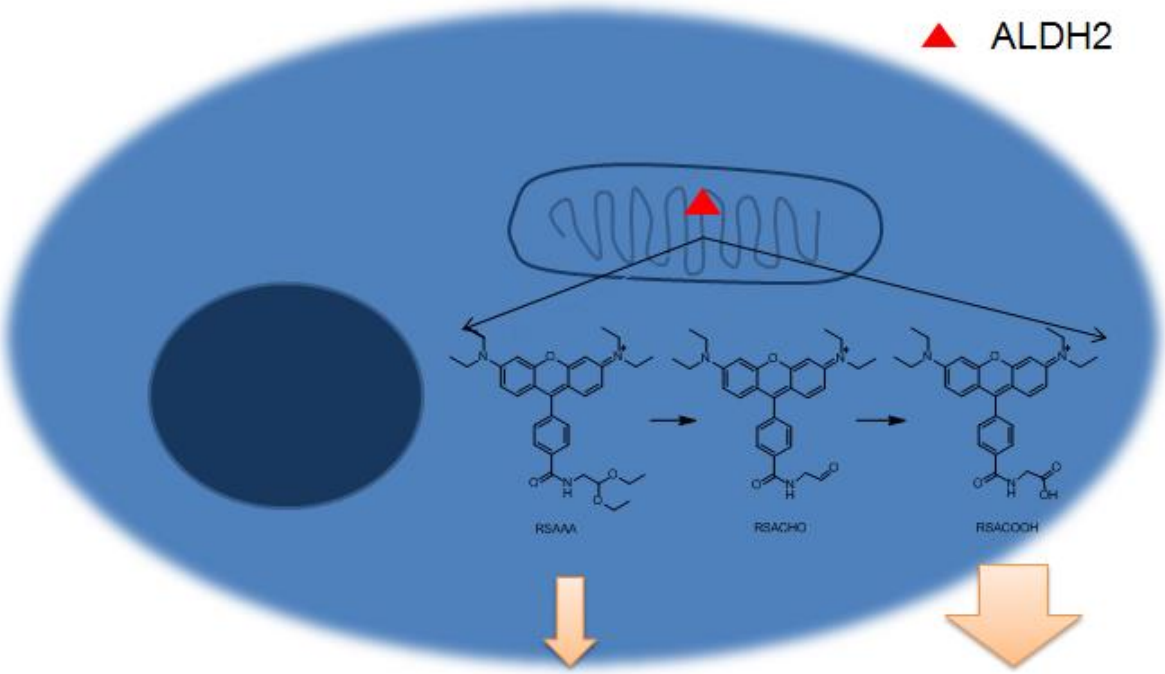


Figure 4.4. Schematic mechanism of assay for ALDH2 activity. Cells expressing high ALDH2 uptook RSAAA and converted it to RSACOOH. RSACOOH was then pumped out of cells, which made the subset of ALDH2-Hi cells to become lower fluorescent. Cold assay buffer was added to prevent the ATP-binding cassette (ABC) transporters to pump out the RSAAA. As a negative control, diethylaminobenzaldehyde (DEAB), a specific ALDH1 and ALDH2 inhibitor, was used to quench the activity of ALDH2-Hi cells, which increased the cells fluorescent signal.

a) b)

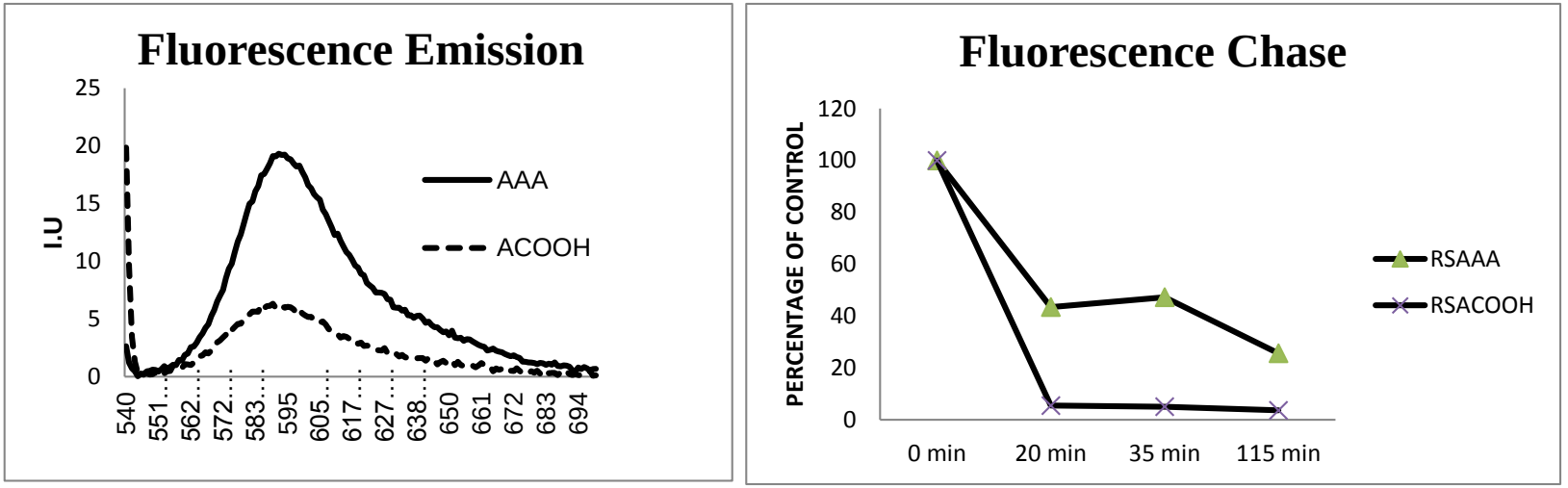


Figure 4.5. Fluorescent property and cell retention of RSAAA and RSACOOH. a) Fluorescence emission file of RSAAA and RSACOOH excited at 536 nm. RSAAA had higher fluorescence intensity than RSACOOH. They both give peak emission at 592 nm. b) Fluorescence Chase experiment of two dyes inside HCT15 cells. Cells were stained with two dyes for 0.5 h. The dyes were then removed and the fluorescence of cells was recorded at different time point. It was found that RSACOOH didn't retain inside cells while more than 30% RSAAA still retain in the cells after 2 h.

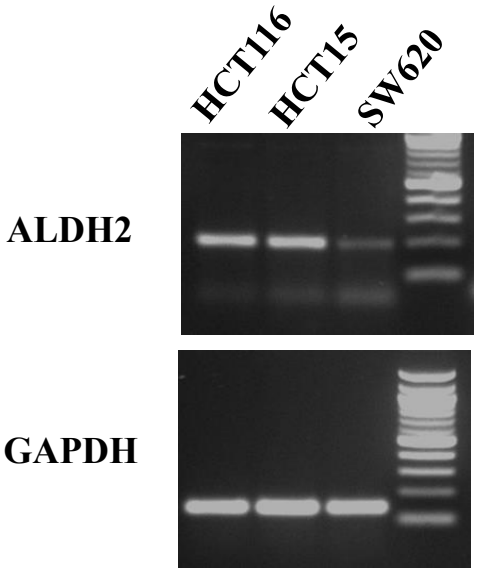
4.2.2 ALDH2 expression and activity in model cell lines

In order to develop the enzyme activity assay in live cell, cell lines with higher expression and activity of ALDH2 were needed. Colon cancers were known to have a various ALDH2 expression [156]. After searching for mRNA database of colon cancer cell lines, we found that the HCT115 might enrich in ALDH2 expression compared to HCT116 and SW620. To check the expression and activity of ALDH2 in these cell lines, Quantitative Reverse Transcription Polymerase Chain Reaction (QRT-PCR) and *in vitro* activity assay were performed. Results from QRT-PCR showed ALDH2 expression in HCT15 cell line was approximately 12 times higher than that in SW620 (Fig. 4.6). ALDH2 activity assay confirmed that there were two fold differences between two cell lines (Fig. 4.7). Therefore, HCT15 was chosen for the following experiment.

4.2.3 Substrate specificity of ALDH2

To test whether RSAAA and RSACOOH might be the potential substrates for ALDH2, the *in vitro* ALDH2 activity assay was then performed using HCT15 cell lysates with these two probes. As showed in Fig. 4.8, RSAAA could be oxidized by cell lysate at a reaction rate around 700 μmol/min/g. The ALDH2 preferred substrate: propionaldehyde was around 2300 μmol/min/g. RSACOOH had a negative reading of reaction rate, which meant it could not be the substrate of ALDH2. The negative reading may be due to the interference from environment. Further control experiments should be performed validate the results.

a)



b)

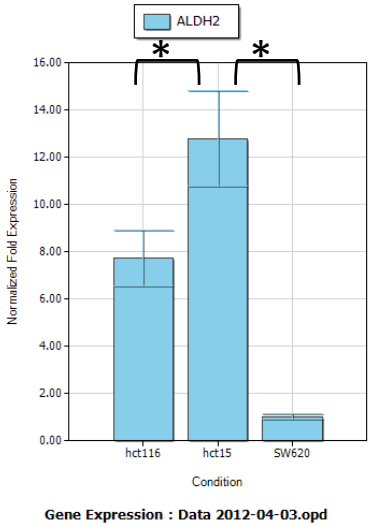
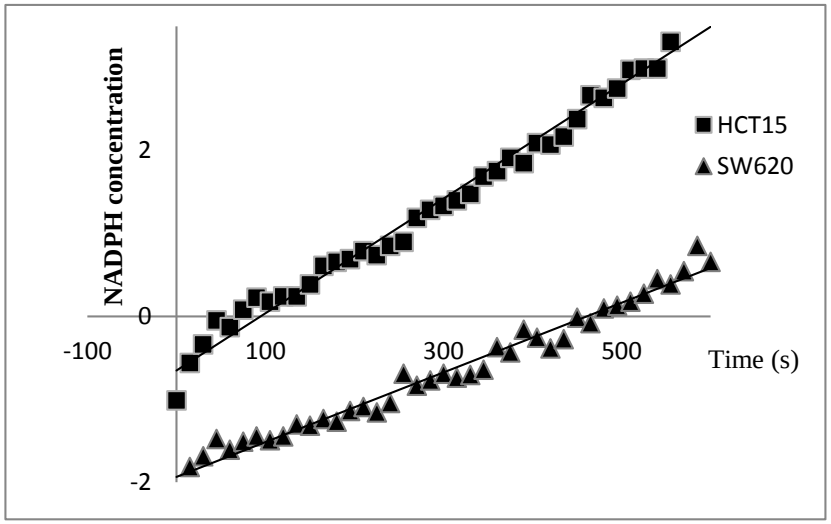


Figure 4.6. Expression of ALDH2 in colon cancer cells. RT-PCR (a) and QRT-PCR (b) were performed on three colon cell lines HCT116, HCT15 and SW620. ALDH2 expression was analyzed using the $2^{-\Delta\Delta CT}$ method [157] for relative quantification. Star indicates significant difference in ALDH2 expression between HCT15 and the other two cell lines by student t test with a 95 confidence interval.HCT15 cells had the highest ALDH2 level.

a)

b)



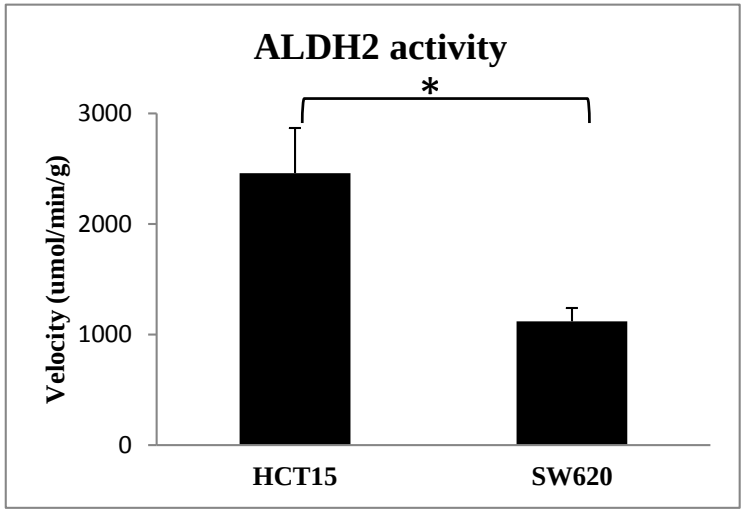


Figure 4.7. Enzyme activity of ALDH2 in different cell lines. Enzyme activity assay of ALDH2 were performed using the cell lysates from HCT15 and SW620. a) The accumulation of substrate over time. b) The ALDH2 activity of two cell line. Star indicated significant difference in ALDH2 activity between HCT15 and SW620 by student t test with a 95 confidence interval. The HCT15 had higher ALDH2 activity than SW620.

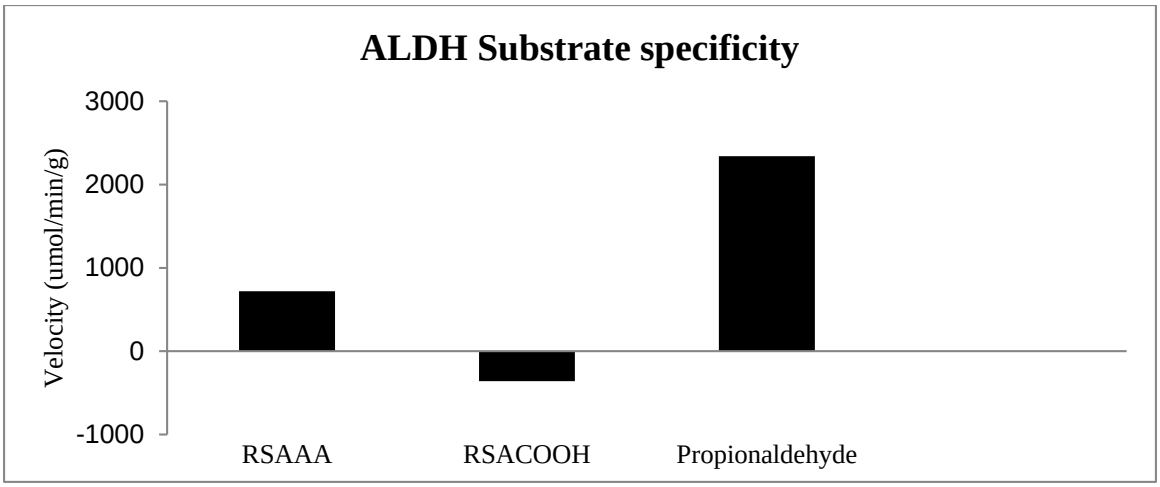


Figure 4.8. Organic probes as ALDH2 substrates. RSAAA, RSACOOH and propionaldehyde were mixed with cell lysates and the ALDH2 reaction velocity were recorded for three substrates. Experiment was only performed once and need to be repeated to further validate the results.

4.2.4 Choice of inhibitor of ALDH2

As showed in Fig. 4.4, to isolate ALDH2 high populations, an inhibitor for ALDH2 was also needed in the design of this protocol. A parallel control experiment with this specific inhibitor would be performed to set up a gate, which controlled the population of inhibitor group to around 1% of total population. This gate would be used to isolate ALDH2 high group in non-inhibitor group, which could be around 10% of the total population according to Aldefluo Assay. Three ALDH2 inhibitors reported in previous studies [137, 158, 159] were screened using *in vitro* activity assay (Fig. 4.9). The working concentrations of inhibitors were chosen according to the reference [137, 158, 159]. It was found that 15 μM DEAB could inhibit the activity as low as 35.4% of the control. Cyanamide could also inhibit ALDH2 activity to 61.4% of wild type but not as good as DEAB. Daidzin could not block ALDH2 activity in our a. It might be due to cell line specificity. DEAB were chosen for its best inhibiting effect on ALDH2 activity.

4.2.5 Localization of fluorescent probes

To test the localization of RSAAA and RSACOOH, a co-staining experiment was performed using LysoTracker or MitoTracker with these two dyes. Both RSAAA and RSACOOH came into the cells and stained the specific subcellular loci. Overlay of two staining channels showed that staining from both dyes were co-localized with MitoTracker (Fig. 4.10) but not with LysoTracker (Fig. 4.11). The results from staining was consistent with our initial design that these probes could preferentially target mitochondria.

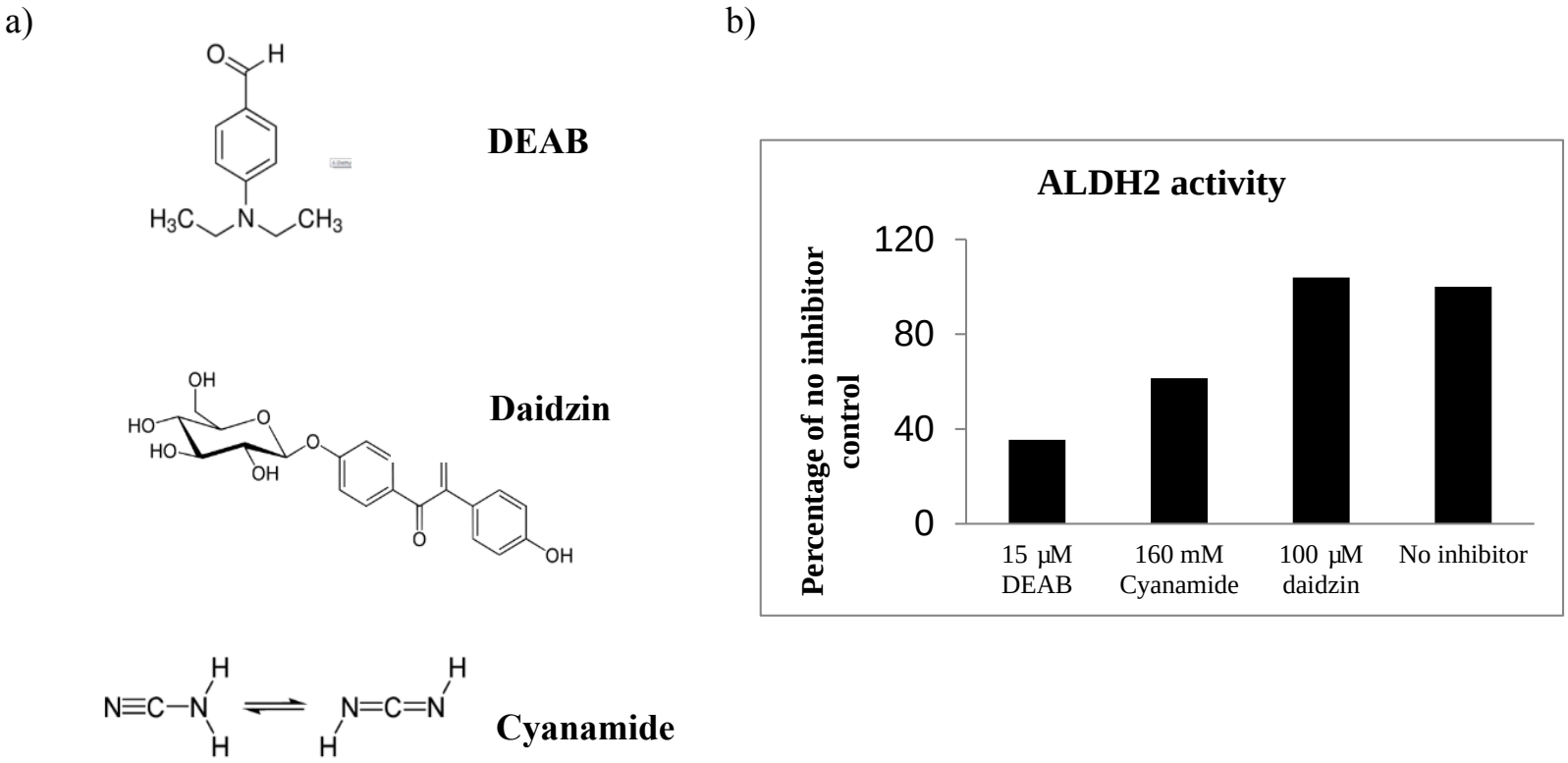


Figure 4.9. Inhibitors of ALDH2. (a) Structures of different inhibitors. (b) The ALDH2 inhibition effect of different inhibitors. The concentration of inhibitors were chosen according to the previous research [137, 158, 159]. Experiment was only performed once and further repeating experiments were needed for statistically analysis.

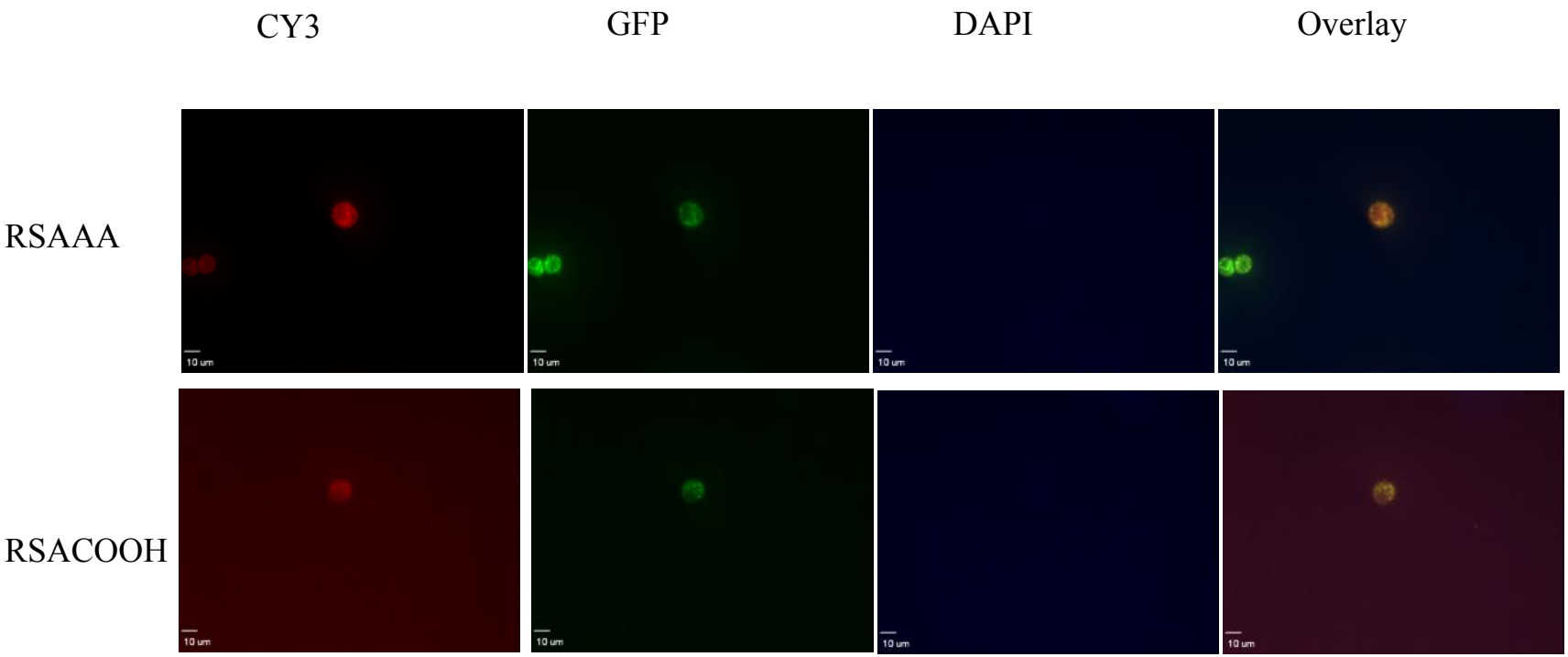


Figure 4.10. Cell co-staining of organic probes and MitoTracker. Cells were stained with either RSAAA or RSACOOH for 0.5 h in the presence of MitoTracker. Cy3 channel tracked organic probes while GFP showed the staining from MitoTracker. DAPI channel was for dead cells staining. Overlay of three channels indicated both RSAAA and RSACOOH co-localized well with MitoTracker. Scale bar: 10 μm.

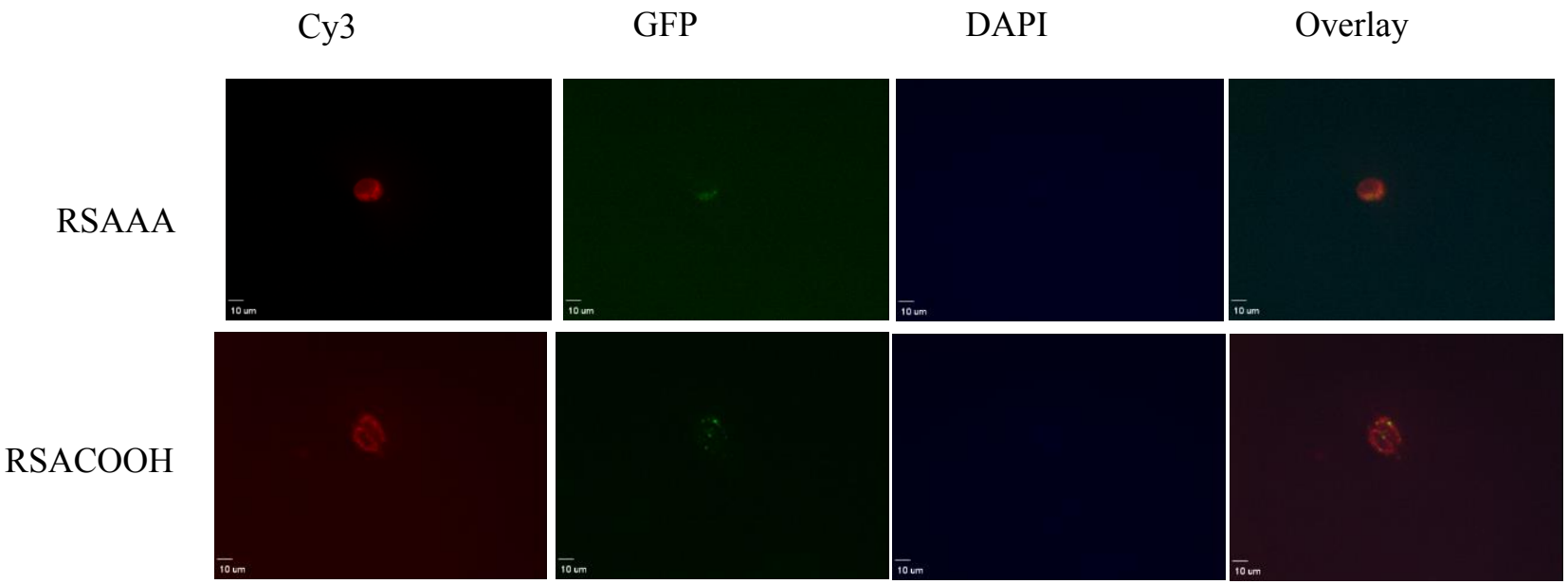


Figure 4.11. Cell co-staining of organic probes and LysoTracker. Cells were stained with either RSAAA or RSACOOH for 0.5 h in the presence of LysoTracker. Cy3 channel tracked staining of organic probes. GFP showed the staining from LysoTracker. DAPI channel were was for dead cells staining. Overlay of three channels indicated both RSAAA and RSACOOH staining not co-localized well with LysoTracker.

4.2.6 DEAB change the fluorescence of cells stained with RSAAA

In Aldefluor Assay, inhibitor treatment should lead to the decrease of fluorescent intensity in cell population which possessed higher ALDH activity. To test whether the designed assay could change the fluorescence, flow cytometry was performed on -/+ DEAB group stained with either RSAAA or RSACOOH. As we expected, treatment with 15 μM DEAB (Fig. 4.12) increased the fluorescence intensity of RSAAA group (from 5141 to 8980) but not RSACOOH (from 482 to 473). This was because DEAB inhibited the ALDH2 activity in cells, which led to the accumulation of RSAAA inside cells. This experiment further confirmed the hypothesis that RSAAA was oxidized by ALDH2 and was converted to RSACOOH, which could be rapidly pumped out of cells.

4.2.7 Isolation of population with different fluorescent intensity and assays for their tumorigenicity

After the selections of substrate and inhibitor, we planned to use this method (we named it AAA assay after the name of the fluorescence probe RSAAA) to isolate ALDH2 high population. In order to mimic Aldefluor assay, two parallel experiments -/+ DEAB were performed. Two gates were set up according to the flow cytometry of + DEAB group in the FACS assay (Fig. 4.13) to isolate three subsets: “Fluo Low”, “Fluo Medium” and “Fluo High” in the -DEAB group. “Fluo Low” was about 1% of the total population with the lowest fluorescence. “Fluo High” was about 5-7% of the total population with the highest fluorescence. “Fluo Medium” was chosen for the middle 40% of “Fluo Low” and “Fluo High”.

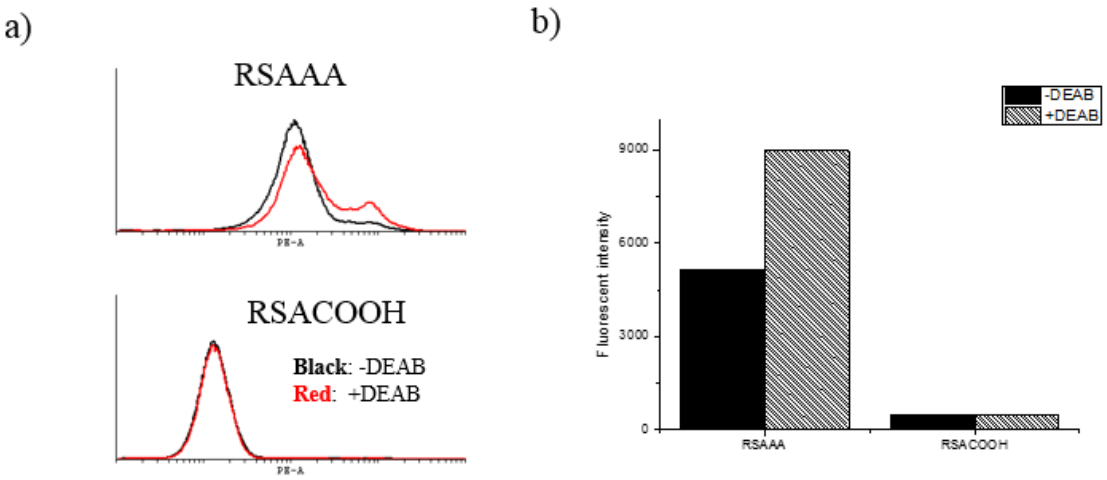


Figure 4.12. Fluorescent intensity changed upon DEAB treatment. a) Cells were treated with or without DEAB and then stained separately either with RSAAA or with RSACOOH. b) Quantification of fluorescence in figure a). Experiment was only performed once. Further repeated experiments were needed for statistically analysis.

+DEAB **-DEAB**

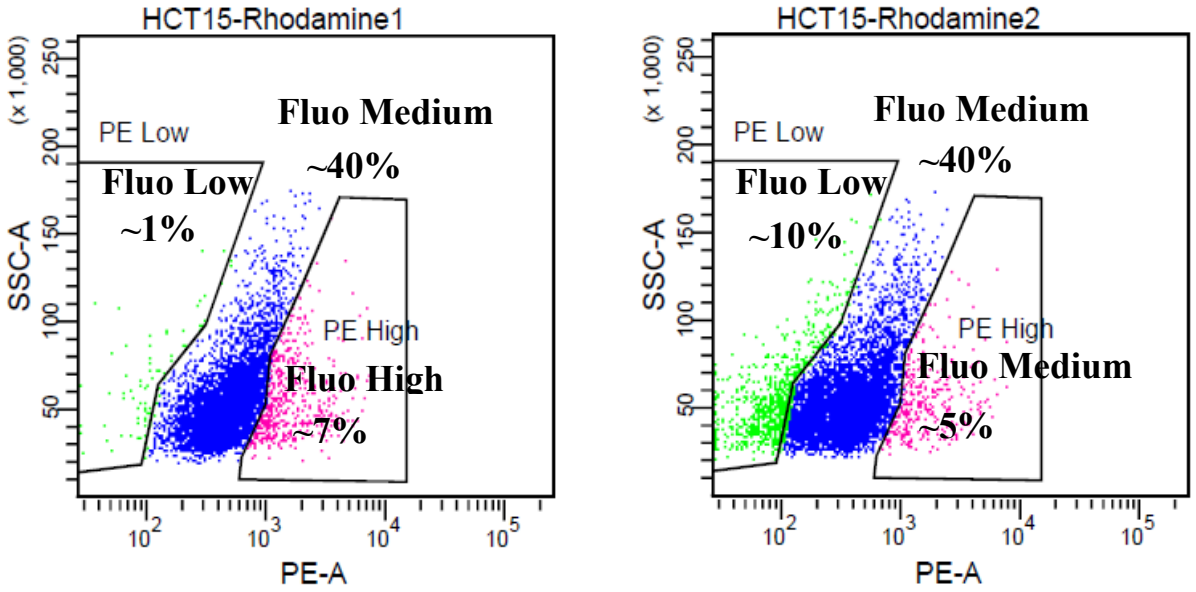


Figure 4.13. Gating of "Fluo Low", "Fluo Medium" and "Fluo High" population. "Fluo Low" was around 1% of total population with the lowest fluorescence signals for DEAB positive group. "Fluo High" was around 5-7% of total population with the highest fluorescence signals for DEAB positive group. "Fluo Medium" was around 10% of total population between "Fluo Low" and "Fluo High" for DEAB positive group. In DEAB negative group, the corresponding "Fluo Low", "Fluo Medium" and "Fluo High" were 10%, 10% and 5% of the total population.

In this case, the cell percentage for three subsets in –DEAB group turned to be around 10%, 40% and 5%. Because the sensitivity of flow cytometry, the total population should not be 100% to optimize the clear isolation of different subsets, especially the "Fluo Low" and "Fluo High" subsets.

To check the tumorigenicity of the "Fluo Low", "Fluo Medium" and "Fluo High" subsets, three different assays were performed to check the potential of being cancer stem cells (Fig. 4.14): sphere assay, soft agar assay and gene expression assay. The first two experiments tested the tumorigenicity potentials in 2D and 3D environment. The gene expression assay by QRT-PCR provided information on important markers indicating cancer stem cells.

We found that "Fluo Low" "Fluo Medium" and "Fluo High" populations each formed 45.3±4.9, 27.5±10.6 and 32.7±4.6 spheres, respectively. For Soft agar assay, they each developed 980±51.3, 1174.7±72.2 and 610±2 colonies. As we can see, "Fluo Low" subsets can form more spheres (Fig. 4.14.a) and colonies (Fig. 4.14.b) than "Fluo High" group. "Fluo Medium" can form more colonies than "Fluo Low" group. No significant difference was found in the potency to form spheres between "Fluo Low" and "Flow Medium" group. "Fluo High" seemed to have lowest potency to form tumor, however the "Fluo Low" and "Fluo Medium" might have comparable tumorigenicity. Statistically analysis were performed using student t test with a 95% confidence interval.

Five genes were chosen for the gene expression analysis, *Nanog* [160], *OCT4* [160, 161], *ALDH1* [140], *Dclk1* [162] and *ALDH2*. The first four were reported as markers for colon cancer stem cells.

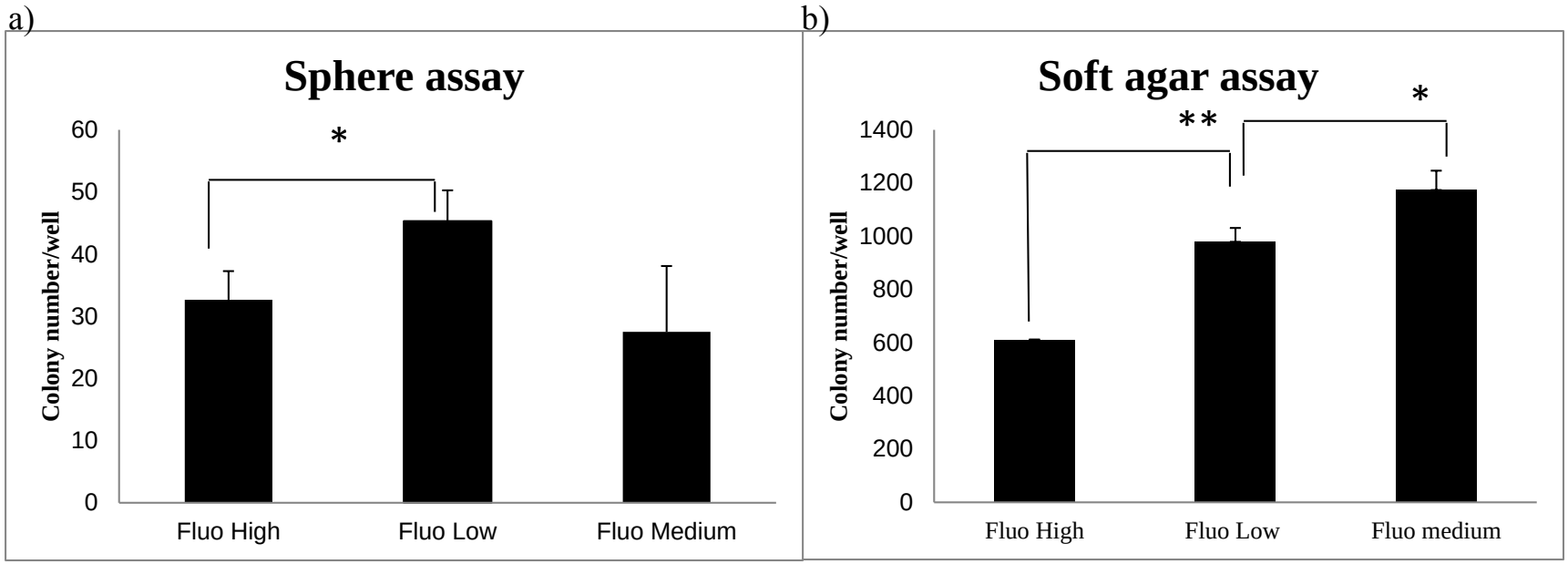


Figure 4.14. Tumorgenicity of AAA sorting populations. a) Sphere assay and b) Soft agar assay were performed to those populations gated by DEAB treatment. Star indicated significant difference by student t test with a 95% confidence interval. Double star indicated the p-value in student t-test is less than 0.001. “Fluo Low” population had higher number of spheres and colonies compared to “Fluo High”. ALDH2 activity might be correlated with tumorgenecity of cell lines.

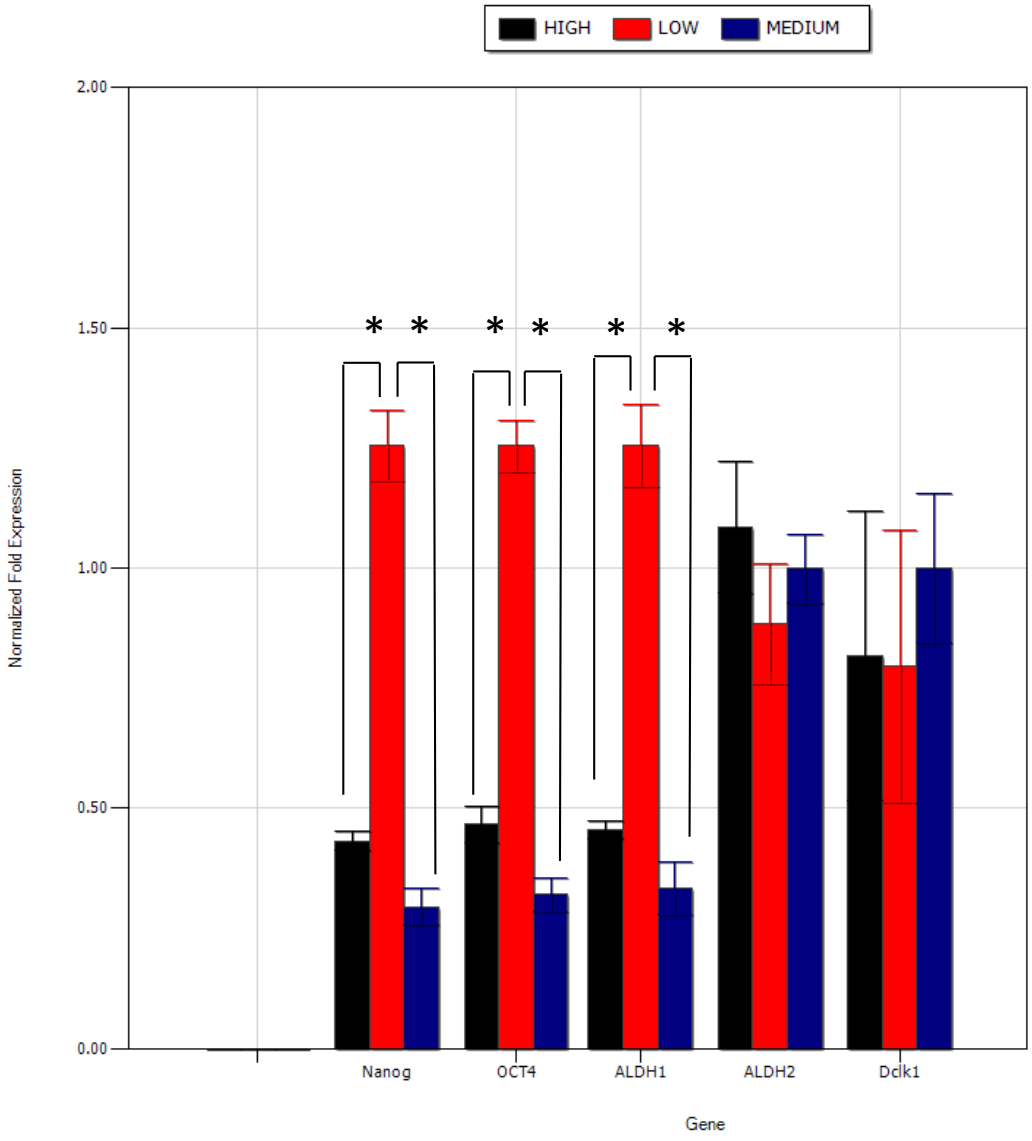


Figure 4.15. Cancer stem cell marker expression in AAA assay sorting populations. Four different cancer stem cells markers' expressions were analyzed using the $2^{-\Delta\Delta CT}$ method [157] for relative quantification. Star indicated significant difference by student t test with a 95% confidence interval.

Three out of four cancer stem cell marker genes expressions were significantly higher in “Fluo low” group: *Nanog, OCT4* and *ALDH1* (Fig. 4.15) by student t-test with a 95% confidence interval. The colon cancer stem cell marker *Dclk1* did not show significantly difference among three populations. Interestingly, no differences were observed for ALDH2 expression. Thus the difference in ALDH2 activity might be due to different activation state of expressed ALDH2 proteins. From the results of the three tumorigenicity experiments, “Fluo Low”, which indicated “high ALDH2 activity” group, might have a higher proportion of cancer stem cells compared to the “Fluo High” group that was the cells with “low ALDH2 activity”. We could use this method to isolate cancer stem cells that had a higher ALDH2 activity in colon cancer cells.

4.3 Conclusion

In order to isolate populations with higher ALDH2 activity in colon cancer, we have synthesized specific fluorescent probes and designed a protocol mimicking Aldefluor Assay. With a positive charge on the backbone, the fluorescent probes RSAAA could be directed to mitochondrion specifically [163]. Since ALDH2 is the major member of ALDH family in mitochondria, the RSAAA could be oxidized and converted to RSACOOH by ALDH2. As the oxidative product of RSAAA, RSACOOH was also synthesized as a staining control. The retentions of RSAAA and RSACOOH were not the same inside cells. Fluorescence chase experiments confirmed RSACOOH was pumped out of cells immediately (less than 20 minutes), while RSAAA staining kept at least 30% fluorescence for up to 2 hours. This

experiment supported the hypothesis that once oxidized by ALDH2, RSAAA could be converted to RSACOOH and pumped out of cells.

These two organic dyes were then tested for their potentials as substrates for ALDH2. Only RSAAA could be oxidized by ALDH2 in the *in vitro* activity assay. Cell imaging from confocal microscope confirmed the localizations of RSAAA and RSACOOH were mainly in mitochondria. Under the treatment of ALDH2 inhibitor, DEAB, the fluorescence inside cells significantly increased, which further supported our hypothesis. Using DEAB treatment to gate for "Fluo Low", " Flow Medium" and "Flow High" population, cell populations with different fluorescent signals were isolated from DEAB negative group. Tumorigenicity experiments on those three populations showed ALDH2 high group ("Fluo low" group) had a higher potential for tumor formation compared to the other two groups.

In conclusion, an assay to assess ALDH2 activity in live colon cancer cells was developed using the organic fluorescent probes synthesized in our lab. Higher ALDH2 activity might be correlated with higher tumorigenicity. This assay might be applied to the isolation of cancer stem cells.

4.4 Experimental Section

4.4.1 Cell maintenance

Please refer to Chapter 1 for HCT116 and HCT15 maintenance. The SW620 cell line was a kindly gift from Prof. Franklin G. Berger (University of South Carolina, Center for Colon Cancer Research). SW620 cells were maintained in RPMI1640 medium

with 10% FBS. All cells were in a water-saturated atmosphere of 5% CO2 and 95% air at 37 ℃.

4.4.2 ALDH2 activity assay

Around 5×10^6 cells were trypsinized and washed twice with PBS. Cells were lysed with 1% Triton X-100 in potassium phosphate buffer [50 mM (pH 7.4)] containing a complete EDTA-free protease inhibitor cocktail (Roche Diagnositic). Cell debris was removed by centrifugation and protein content was determined with the Commassie Blue Kit from Thermo Scientific. The reaction mixture contained 40 mM (pH 7.4) potassium phosphate buffer, 4 mM dithiothreitol, 5 mM $MgCl_2$, 10 mM 4-methyl-pyrazole, 0.8 mM β-nicotinamide adenine dinucleotide (β-NAD+) and 100 μg of total protein. The reaction was initiated with the addition of 14 μM propionaldehyde and production of NADH was measured at 340 nm over time. The specific activity was expressed as nmol of NADH/min/mg of protein.

4.4.3 Cell imaging

Confocal fluorescence imaging was performed with an Olympus X81 fluorescence microscope. Cells were incubated with sensors and trackers in the presence of (+/-)-Verapamil Hydrochloride (from MP Biomedicals, Cat. No: 195545) for 0.5 hour at 37 ℃. Fluorescence images of cells were taken after washed with PBS buffer.

4.4.4 Flow cytometry

Cellular ALDH2 activity was quantified by flow cytometric analysis. HCT15 cells were seed on 10 cm plates to allow adhere and grow overnight. Cells were trypsinized the next day and 1×10^6 cells were collected for each sample. After washed once with PBS, cells were stained with indicated dyes for 0.5 hour at 37 ℃. The samples were washed once with cold PBS and resuspended in PBS + 0.1% FBS. Samples were analyzed by BD LSRII flow cytometer (BD Bioscience). For DEAB treatment group, DEAB was in the staining buffer and washing buffers for each step.

4.4.5 Cell sorting

To get enough populations, 2×10^7 cells were collected for each sample and subjected to sort by BD FACS Canto I (BD Bioscinece). Cells were sorted according to the gates set up by DEAB treatment group. Sorted cells were kept in full medium on ice for the next steps.

4.4.6 Mammal sphere assay

Sorted cells were washed once with PBS and resuspended in mammal sphere culture medium (Serum-free DMEM/F12 (Invitrogen) and supplemented with 20 ng/mL epidermal growth factor (Stemgent, Inc.), 10 ng/mL basic fibroblast growth factor (Stemgent, Inc.), 10 μg/mL insulin (Sigma), 1× B27 supplement (Invitrogen), 100 U/mL penicillin, 100 μg/mL streptomycin and 0.5% bovine serum albumin (Sigma)). Cells were seeded in 1% agarose pre-coated 12 well plate at density of 10000 cells/mL. Cells were cultured in 5% CO2 at 37 ℃ for a week. Spheres of more than 20 cells were counted

under Olympus X51 microscope. Each experiment was performed in triplicate. Error bars represented ± standard deviation.

4.4.7 Soft agar assay

Sorted cells were washed with PBS once and 5000 cells from each sample were seeded in one well of 6-well plate in 0.3% low melting point agarose (Invitrogen) with culture medium on top of solidified 0.5% low-melting point agarose with culture medium. Cells were cultured in CO_2 and 95% air at 37 ℃ for two weeks and 1 mL of 0.005% crystal violet was added to visualize the colonies before counting. Colonies were counted by naked eye. Images of soft agar assay were obtained by gel imaging system GelDoc XR (Bio-Rad). Each experiment was performed in triplicate. Error bars represent ± standard deviation.

4.4.7 Gene expression analysis

QRT-PCR was used to evaluate gene expression level by measuring relative quantity of mRNA of each gene. Total RNA was extracted from the cells using RNeasy mini purification kit according to manufacturer's protocol (Qiagen), and subsequently reverse-transcribed with a qSCript cDNA synthesis kit (Quanta Biosciences). RT-qPCR (iQ5 real-time PCR detection system, Bio-Rad) was done by the method described as: 45 cycles of PCR (95 ℃ for 15 s, 60 ℃ for 15 s, and 72 ℃ for 30s), after initial denaturation step of 5 minutes at 95 °C , by using 12.5 μL of iQ5 SYBR Green I supermix, 2 pmol/μL of each forward and reverse primers and 0.5 μL cDNA templates in a final reaction volume of 25 μL. Data collection

was enabled at 72 °C in each cycle and C_T (threshold cycle) values were calculated using the iQ5 optical system software version 2.1. The primers used for RT-qPCR are showed in Table 4.2. The primers were reported previously in literatures [160] or designed by our lab. Data analysis was performed using the $2^{-\Delta\Delta CT}$ method [157] for relative quantification. The primers were then synthesized commercially (Integrated DNA Technologies, Inc.), and evaluated for an annealing temperature of 60˚C. All samples were normalized to Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) as the house keeping gene. Each experiment was performed in triplicate. Error bars represent ± standard deviation.

Table 4.2. Primers of cancer stem cell markers

Primer name	Gene name	Specie	Sequence
OCT4 F	OCT4	HUMAN	CCGTGAAGCTGGAGAAGGAGAAG
OCT4 R	OCT4	HUMAN	AGCGGCAGATGGTCGTTTGG
Nanog F	NANOG	HUMAN	GAACTCTCCAACATCCTGAACCTC
Nanog R	NANOG	HUMAN	CCTTCTGCGTCACACCATTGC
ALDH1 F	ALDH1	HUMAN	AGCAGGAGTGTTTACCAAAGA
ALDH1 R	ALDH1	HUMAN	CCCAGTTCTCTTCCATTTCCAG
ALDH2 F	ALDH2	HUMAN	ACAAGGAAGATGTGGACAAGG
ALDH2 R	ALDH2	HUMAN	ATGGTCCTCAAATGTCTCCG
Dclk1 F	Dclk1	HUMAN	TGATGCTGTTGGTCGATGTAG
Dclk1 R	Dclk1	HUMAN	AGTGCTATGACAGAAACTCCAG

REFERENCES

[1] Zhang S, Ong C-N, Shen H-M. Critical roles of intracellular thiols and calcium in parthenolide-induced apoptosis in human colorectal cancer cells. Cancer Letters. 2004;208:143-53.

[2] Stipanuk MH, Dominy JE, Lee J-I, Coloso RM. Mammalian cysteine metabolism: new insights into regulation of cysteine metabolism. The Journal of Nutrition. 2006;136:1652S-9S.

[3] El-Khairy L, Ueland PM, Refsum H, Graham IM, Vollset SE. Plasma total cysteine as a risk factor for vascular disease The European Concerted Action Project. Circulation. 2001;103:2544-9.

[4] Tolar J, Orchard P, Bjoraker K, Ziegler R, Shapiro E, Charnas L. N-acetyl-L-cysteine improves outcome of advanced cerebral adrenoleukodystrophy. Bone marrow transplantation. 2007;39:211-5.

[5] Park S, Imlay JA. High levels of intracellular cysteine promote oxidative DNA damage by driving the fenton reaction. Journal of bacteriology. 2003;185:1942-50.

[6] Bradley H, Gough A, Sokhi R, Hassell A, Waring R, Emery P. Sulfate metabolism is abnormal in patients with rheumatoid arthritis. Confirmation by in vivo biochemical findings. The Journal of Rheumatology. 1994;21:1192-6.

[7] Jameson GN. Iron, cysteine and Parkinson's disease. Monatshefte für Chemie-Chemical Monthly. 2011;142:325-9.

[8] Heafield MT, Fearn S, Steventon GB, Waring RH, Williams AC, Sturman SG. Plasma cysteine and sulphate levels in patients with motor neurone, Parkinson's and Alzheimer's disease. Neuroscience Letters. 1990;110:216-20.

[9] Persichilli S, Gervasoni J, Castagnola M, Zuppi C, Zappacosta B. A reversed-phase HPLC fluorimetric method for simultaneous determination of homocysteine-related thiols in different body fluids. LabMed. 2011;42:657-62.

[10] Wald DS, Law M, Morris JK. Homocysteine and cardiovascular disease: evidence on causality from a meta-analysis. British Medical Journal. 2002;325:1202.

[11] Dominy JE, Hwang J, Stipanuk MH. Overexpression of cysteine dioxygenase reduces intracellular cysteine and glutathione pools in HepG2/C3A cells. American Journal of Physiology-Endocrinology and Metabolism. 2007;293:E62-E9.

[12] Ravaglia G, Forti P, Maioli F, Martelli M, Servadei L, Brunetti N, et al. Homocysteine and folate as risk factors for dementia and Alzheimer disease. American Journal of clinical Nutrition. 2005;82:636-43.

[13] Seshadri S, Beiser A, Selhub J, Jacques PF, Rosenberg IH, D'Agostino RB, et al. Plasma homocysteine as a risk factor for dementia and Alzheimer's disease. New England Journal of Medicine. 2002;346:476-83.

[14] Hong R, Han G, Fernández JM, Kim B-j, Forbes NS, Rotello VM. Glutathione-mediated delivery and release using monolayer protected nanoparticle carriers. Journal of the American Chemical Society. 2006;128:1078-9.

[15] Hassan SS, Rechnitz G. Determination of glutathione and glutathione reductase with a silver sulfide membrane electrode. Analytical Chemistry. 1982;54:1972-6.

[16] Schafer FQ, Buettner GR. Redox environment of the cell as viewed through the redox state of the glutathione disulfide/glutathione couple. Free Radical Biology and Medicine. 2001;30:1191-212.

[17] Filomeni G, Rotilio G, Ciriolo MR. Cell signalling and the glutathione redox system. Biochemical Pharmacology. 2002;64:1057-64.

[18] Fratelli M, Goodwin LO, Ørom UA, Lombardi S, Tonelli R, Mengozzi M, et al. Gene expression profiling reveals a signaling role of glutathione in redox regulation. Proceedings of the National Academy of Sciences USA. 2005;102:13998-4003.

[19] Dalton TP, Shertzer HG, Puga A. Regulation of gene expression by reactive oxygen. Annual Review of Pharmacology and Toxicology. 1999;39:67-101.

[20] Townsend DM, Tew KD. The role of glutathione-S-transferase in anti-cancer drug resistance. Oncogene. 2003;22:7369-75.

[21] Aksenov MY, Markesbery WR. Changes in thiol content and expression of glutathione redox system genes in the hippocampus and cerebellum in Alzheimer's disease. Neuroscience Letters. 2001;302:141-5.

[22] Kleinman WA, Richie Jr JP. Status of glutathione and other thiols and disulfides in human plasma. Biochemical Pharmacology. 2000;60:19-29.

[23] Blankenberg S, Rupprecht HJ, Bickel C, Torzewski M, Hafner G, Tiret L, et al. Glutathione peroxidase 1 activity and cardiovascular events in patients with coronary artery disease. New England Journal of Medicine. 2003;349:1605-13.

[24] Finkelstein J. The metabolism of homocysteine: pathways and regulation. European Journal of Pediatrics. 1998;157:S40-S4.

[25] Stipanuk MH, Coloso RM, Garcia R, Banks MF. Cysteine concentration regulates cysteine metabolism to glutathione, sulfate and taurine in rat hepatocytes. The Journal of Nutrition. 1992;122:420-7.

[26] Wu G, Fang Y-Z, Yang S, Lupton JR, Turner ND. Glutathione metabolism and its implications for health. The Journal of Nutrition. 2004;134:489-92.

[27] De Rosa S, Zaretsky M, Dubs J, Roederer M, Anderson M, Green A, et al. N-acetylcysteine replenishes glutathione in HIV infection. European Journal of Clinical Investigation. 2000;30:915-29.

[28] Arakawa M, Ito Y. N-acetylcysteine and neurodegenerative diseases: basic and clinical pharmacology. The Cerebellum. 2007;6:308-14.

[29] De Flora S, Izzotti A, D'Agostini F, Balansky RM. Mechanisms of N-acetylcysteine in the prevention of DNA damage and cancer, with special reference to smoking-related end-points. Carcinogenesis. 2001;22:999-1013.

[30] Zhang W, Trachootham D, Liu J, Chen G, Pelicano H, Garcia-Prieto C, et al. Stromal control of cystine metabolism promotes cancer cell survival in chronic lymphocytic leukaemia. Nature Cell Biology. 2013;14:276-86.

[31] Laxman S, Sutter BM, Wu X, Kumar S, Guo X, Trudgian DC, et al. Sulfur amino acids regulate translational capacity and metabolic homeostasis through modulation of tRNA thiolation. Cell. 2013;154:416-29.

[32] Huang Y, Dai Z, Barbacioru C, Sadee W. Cystine-glutamate transporter SLC7A11 in cancer chemosensitivity and chemoresistance. Cancer Research. 2005;65:7446-54.

[33] Chen Y, Swanson RA. The glutamate transporters EAAT2 and EAAT3 mediate cysteine uptake in cortical neuron cultures. Journal of Neurochemistry. 2003;84:1332-9.

[34] Conrad M, Sato H. The oxidative stress-inducible cystine/glutamate antiporter, system x c−: cystine supplier and beyond. Amino Acids. 2012;42:231-46.

[35] Tietze F, Bradley K, Schulman J. Enzymic Reduction of Cystine by Subcellular Fractions of Cultured and Peripheral Leukoyctes from Normal and Cystinotic Individuals. Pediatric Research. 1972;6:649-58.

[36] Berman AE, Chan WY, Brennan AM, Reyes RC, Adler BL, Suh SW, et al. N‐acetylcysteine prevents loss of dopaminergic neurons in the EAAC1−/− mouse. Annals of Neurology. 2011;69:509-20.

[37] Paul BD, Sbodio JI, Xu R, Vandiver MS, Cha JY, Snowman AM, et al. Cystathionine gamma-lyase deficiency mediates neurodegeneration in Huntington's disease. Nature. 2014;509:96-100.

[38] Martinez-Banaclocha MA. N-acetyl-cysteine in the treatment of Parkinson's disease. What are we waiting for? Medical Hypotheses. 2012;79:8-12.

[39] Fiskerstrand T, Refsum H, Kvalheim G, Ueland PM. Homocysteine and other thiols in plasma and urine: automated determination and sample stability. Clinical Chemistry. 1993;39:263-71.

[40] Stabler SP, Marcell PD, Podell ER, Allen RH. Quantitation of total homocysteine, total cysteine, and methionine in normal serum and urine using capillary gas chromatography-mass spectrometry. Analytical Biochemistry. 1987;162:185-96.

[41] Inoue T, Kirchhoff JR. Determination of thiols by capillary electrophoresis with amperometric detection at a coenzyme pyrroloquinoline quinone modified electrode. Analytical Chemistry. 2002;74:1349-54.

[42] Jung HS, Chen X, Kim JS, Yoon J. Recent progress in luminescent and colorimetric chemosensors for detection of thiols. Chemical Society Reviews. 2013;42:6019-31.

[43] Snyder LR, Kirkland JJ, Glajch JL. Practical HPLC method development: John Wiley & Sons; 2012.

[44] Tanaka F, Mase N, Barbas Iii CF. Determination of cysteine concentration by fluorescence increase: reaction of cysteine with a fluorogenic aldehyde. Chemical Communications. 2004:1762-3.

[45] Zhang M, Yu M, Li F, Zhu M, Li M, Gao Y, et al. A highly selective fluorescence turn-on sensor for cysteine/homocysteine and its application in bioimaging. Journal of the American Chemical Society. 2007;129:10322-3.

[46] Li H, Liu F, Xiao Y, Pellechia PJ, Smith MD, Qian X, et al. Revisit of a series of ICT fluorophores: skeletal characterization, structural modification, and spectroscopic behavior. Tetrahedron. 2014;70:5872-7.

[47] Jones DP, Go Y-M, Anderson CL, Ziegler TR, KINKADE JM, Kirlin WG. Cysteine/cystine couple is a newly recognized node in the circuitry for biologic redox signaling and control. The FASEB Journal. 2004;18:1246-8.

[48] Siegel R, Naishadham D, Jemal A. Cancer statistics, 2013. CA-A Cancer Journal for Clinicians. 2013;63:11-30.

[49] Faleiro-Rodrigues C, Macedo-Pinto I, Pereira D, Lopes CS. Prognostic value of E-cadherin immunoexpression in patients with primary ovarian carcinomas. Annals of Oncology. 2004;15:1535-42.

[50] Burstein HJ. The distinctive nature of HER2-positive breast cancers. New England Journal of Medicine. 2005;353:1652-4.

[51] Mitri Z, Constantine T, O'Regan R. The HER2 receptor in breast cancer: pathophysiology, clinical use, and new advances in therapy. Chemotherapy research and practice. 2012;2012:743193.

[52] Coussens L, Yang-Feng TL, Liao Y-C, Chen E, Gray A, McGrath J, et al. Tyrosine kinase receptor with extensive homology to EGF receptor shares chromosomal location with neu oncogene. Science. 1985;230:1132-9.

[53] Herbst RS. Review of epidermal growth factor receptor biology. International Journal of Radiation Oncology Biology Physics. 2004;59:S21-S6.

[54] Herbst RS. Review of epidermal growth factor receptor biology. International Journal of Radiation Oncology* Biology* Physics. 2004;59:S21-S6.

[55] Olayioye MA. Update on HER-2 as a target for cancer therapy: intracellular signaling pathways of ErbB2/HER-2 and family members. Breast Cancer Research. 2001;3:385-9.

[56] Gasol E, Jiménez-Vidal M, Chillarón J, Zorzano A, Palacín M. Membrane topology of system xc-light subunit reveals a re-entrant loop with substrate-restricted accessibility. Journal of Biological Chemistry. 2004;279:31228-36.

[57] Liu X, Li X, Zhang B, Liang Y, Zhou C, Cao D, et al. MicroRNA-26b is underexpressed in human breast cancer and induces cell apoptosis by targeting SLC7A11. FEBS Letters. 2011;585:1363-7.

[58] Wang SE, Narasanna A, Perez-Torres M, Xiang B, Wu FY, Yang S, et al. HER2 kinase domain mutation results in constitutive phosphorylation and activation of HER2 and EGFR and resistance to EGFR tyrosine kinase inhibitors. Cancer Cell. 2006;10:25-38.

[59] Armstrong JS, Whiteman M, Yang H, Jones DP, Sternberg P, Jr. Cysteine starvation activates the redox-dependent mitochondrial permeability transition in retinal pigment epithelial cells. Investigative ophthalmology & visual science. 2004;45:4183-9.

[60] Kwon YH, Stipanuk MH. Cysteine regulates expression of cysteine dioxygenase and gamma-glutamylcysteine synthetase in cultured rat hepatocytes. American Journal of Physiology-Endocrinology and Metabolism. 2001;280:E804-E15.

[61] Chen Z, Wang X, Zhu W, Cao X, Tong L, Li H, et al. Acenaphtho [1, 2-b] pyrrole-Based selective fibroblast growth factor receptors 1 (FGFR1) Inhibitors: design, synthesis, and biological Activity. Journal of Medicinal Chemistry. 2011;54:3732-45.

[62] Song T, Li X, Chang X, Liang X, Zhao Y, Wu G, et al. 3-Thiomorpholin-8-oxo-8*H*-acenaphtho [1, 2-*b*] pyrrole-9-carbonitrile (S1) derivatives as pan-Bcl-2-inhibitors of Bcl-2, Bcl-$x_L$and Mcl-1. Bioorganic & Medicinal Chemistry. 2013;21:11-20.

[63] Zhang Z, Wu G, Xie F, Song T, Chang X. 3-Thiomorpholin-8-oxo-8 H-acenaphtho [1, 2-b] pyrrole-9-carbonitrile (S1) based molecules as potent, dual inhibitors of B-cell lymphoma 2 (Bcl-2) and myeloid cell leukemia sequence 1 (Mcl-1): structure-based design and structure− activity relationship studies. Journal of Medicinal Chemistry. 2011;54:1101-5.

[64] Spector N, Xia W, El-Hariry I, Yarden Y, Bacus S. HER2 therapy. Small molecule HER-2 tyrosine kinase inhibitors. Breast Cancer Research. 2007;9:205.

[65] Atkinson AJ, Huang S-M, Markey SP. Principles of clinical pharmacology: Academic Press; 2012.

[66] Boyd MR. The NCI in vitro anticancer drug discovery screen. Anticancer Drug Development Guide: Springer; 1997. p. 23-42.

[67] Ghavami S, Hashemi M, Ande SR, Yeganeh B, Xiao W, Eshraghi M, et al. Apoptosis and cancer: mutations within caspase genes. Journal of Medical Genetics. 2009;46:497-510.

[68] Kasai H, Nakanishi K, Macfarlane R, Torgerson D, Ohashi Z, McCloskey J, et al. The structure of Q* nucleoside isolated from rabbit liver transfer ribonucleic acid. Journal of the American Chemical Society. 1976;98:5044-6.

[69] Mitra N, Sharon N, Surolia A. Role of N-linked glycan in the unfolding pathway of Erythrina corallodendron lectin. Biochemistry. 2003;42:12208-16.

[70] Grobe K, Ledin J, Ringvall M, Holmborn K, Forsberg E, Esko JD, et al. Heparan sulfate and development: differential roles of the N-acetylglucosamine N-deacetylase/ N-sulfotransferase isozymes. Biochimica et Biophysica Acta (BBA)-General Subjects. 2002;1573:209-15.

[71] Apweiler R, Hermjakob H, Sharon N. On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. Biochimica et Biophysica Acta (BBA)-General Subjects. 1999;1473:4-8.

[72] Jung E, Veuthey AL, Gasteiger E, Bairoch A. Annotation of glycoproteins in the SWISS‐PROT database. Proteomics. 2001;1:262-8.

[73] Haines N, Irvine KD. Glycosylation regulates Notch signalling. Nature Reviews Molecular Cell Biology. 2003;4:786-97.

[74] Herrmann M, von der Lieth CW, Stehling P, Reutter W, Pawlita M. Consequences of a subtle sialic acid modification on the murine polyomavirus receptor. Journal of Virology. 1997;71:5922-31.

[75] Ahn YH, Shin PM, Kim Y-S, Oh NR, Ji ES, Kim KH, et al. Quantitative analysis of aberrant protein glycosylation in liver cancer plasma by AAL-enrichment and MRM mass spectrometry. Analyst. 2013;138:6454-62.

[76] Tian Y, Zhang H. Glycoproteomics and clinical applications. Proteomics-Clinical Applications. 2010;4:124-32.

[77] Zhu J, Wang Y, Yu Y, Wang Z, Zhu T, Xu X, et al. Aberrant fucosylation of glycosphingolipids in human hepatocellular carcinoma tissues. Liver International. 2014;34:147-60.

[78] Xiong L, Andrews D, Regnier F. Comparative proteomics of glycoproteins based on lectin selection and isotope coding. Journal of Proteome Research. 2003;2:618-25.

[79] Hahne H, Sobotzki N, Nyberg T, Helm D, Borodkin VS, van Aalten DM, et al. Proteome wide purification and identification of O-GlcNAc-modified proteins using click chemistry and mass spectrometry. Journal of Proteome Research. 2013;12:927-36.

[80] Hizal DB, Wolozny D, Colao J, Jacobson E, Tian Y, Krag SS, et al. Glycoproteomic and glycomic databases. Clinical Proteomics. 2014;11:15.

[81] West I, Goldring O. Lectin affinity chromatography. Protein Purification Protocols: Springer; 1996. p. 177-86.

[82] Sunderland CA, McMaster WR, Williams AF. Purification with monoclonal antibody of a predominant leukocyte-common antigen and glycoprotein from rat thymocytes. European Journal of Immunology. 1979;9:155-9.

[83] Slade PG, Hajivandi M, Bartel CM, Gorfien SF. Identifying the CHO secretome using mucin-type O-linked glycosylation and click-chemistry. Journal of Proteome Research. 2012;11:6175-86.

[84] Wada Y, Azadi P, Costello CE, Dell A, Dwek RA, Geyer H, et al. Comparison of the methods for profiling glycoprotein glycans—HUPO Human Disease Glycomics/Proteome Initiative multi-institutional study. Glycobiology. 2007;17:411-22.

[85] Wada Y, Dell A, Haslam SM, Tissot B, Canis K, Azadi P, et al. Comparison of methods for profiling o-glycosylation human proteome organisation human disease glycomics/proteome initiative multi-institutional study of iga1. Molecular & Cellular Proteomics. 2010;9:719-27.

[86] Kolb HC, Finn M, Sharpless KB. Click chemistry: diverse chemical function from a few good reactions. Angewandte Chemie International Edition. 2001;40:2004-21.

[87] Kolb HC, Sharpless KB. The growing impact of click chemistry on drug discovery. Drug Discovery Today. 2003;8:1128-37.

[88] Zaro BW, Hang HC, Pratt MR. Incorporation of unnatural sugars for the identification of glycoproteins. Mass Spectrometry of Glycoproteins: Springer; 2013. p. 57-67.

[89] Shieh P, Siegrist MS, Cullen AJ, Bertozzi CR. Imaging bacterial peptidoglycan with near-infrared fluorogenic azide probes. Proceedings of the National Academy of Sciences. 2014;111:5456-61.

[90] Saxon E, Bertozzi CR. Cell surface engineering by a modified Staudinger reaction. Science. 2000;287:2007-10.

[91] Baskin JM, Prescher JA, Laughlin ST, Agard NJ, Chang PV, Miller IA, et al. Copper-free click chemistry for dynamic in vivo imaging. Proceedings of the National Academy of Sciences. 2007;104:16793-7.

[92] Dube DH, Prescher JA, Quang CN, Bertozzi CR. Probing mucin-type O-linked glycosylation in living animals. Proceedings of the National Academy of Sciences USA. 2006;103:4819-24.

[93] Sletten EM, Bertozzi CR. From mechanism to mouse: a tale of two bioorthogonal reactions. Accounts of Chemical Research. 2011;44:666-76.

[94] Gold B, Shevchenko NE, Bonus N, Dudley GB, Alabugin IV. Selective transition state stabilization via hyperconjugative and conjugative assistance: stereoelectronic concept for copper-free click chemistry. The Journal of Organic Chemistry. 2011;77:75-89.

[95] Agard NJ, Bertozzi CR. Chemical approaches to perturb, profile, and perceive glycans. Accounts of Chemical Research. 2009;42:788-97.

[96] Boyce M, Carrico IS, Ganguli AS, Yu S-H, Hangauer MJ, Hubbard SC, et al. Metabolic cross-talk allows labeling of O-linked β-N-acetylglucosamine-modified proteins via the N-acetylgalactosamine salvage pathway. Proceedings of the National Academy of Sciences. 2011;108:3141-6.

[97] Zhang Y, Yuan J, Song J, Wang Z, Huang L. An efficient method for selectively imaging and quantifying in situ the expression of sialylated glycoproteins on living cells. Glycobiology. 2013;23:643-53.

[98] Laughlin ST, Baskin JM, Amacher SL, Bertozzi CR. In vivo imaging of membrane-associated glycans in developing zebrafish. Science. 2008;320:664-7.

[99] Pan S, Chen R, Tamura Y, Crispin DA, Lai LA, May DH, et al. Quantitative glycoproteomics analysis reveals changes in N-Glycosylation level associated with Pancreatic Ductal Adenocarcinoma. Journal of Proteome Research. 2014;13:1293-306.

[100] Gao X, Yang L, Petros JA, Marshall FF, Simons JW, Nie S. < i> In vivo</i> molecular and cellular imaging with quantum dots. Current Opinion in Biotechnology. 2005;16:63-72.

[101] Jaiswal JK, Goldman ER, Mattoussi H, Simon SM. Use of quantum dots for live cell imaging. Nature Methods. 2004;1:73-8.

[102] Michalet X, Pinaud FF, Bentolila LA, Tsay JM, Doose S, Li JJ, et al. Quantum dots for live cells, in vivo imaging, and diagnostics. Science. 2005;307:538-44.

[103] Medintz IL, Uyeda HT, Goldman ER, Mattoussi H. Quantum dot bioconjugates for imaging, labelling and sensing. Nature Materials. 2005;4:435-46.

[104] Mattoussi H, Mauro JM, Goldman ER, Anderson GP, Sundar VC, Mikulec FV, et al. Self-assembly of CdSe-ZnS quantum dot bioconjugates using an engineered recombinant protein. Journal of the American Chemical Society. 2000;122:12142-50.

[105] Hang HC, Yu C, Kato DL, Bertozzi CR. A metabolic labeling approach toward proteomic analysis of mucin-type O-linked glycosylation. Proceedings of the National Academy of Sciences. 2003;100:14846-51.

[106] Prescher JA, Dube DH, Bertozzi CR. Chemical remodelling of cell surfaces in living animals. Nature. 2004;430:873-7.

[107] Montreuil J, Schachter H, Vliegenthart J. Primary structure of glycoprotein glycans. Glycoproteins I. 1995;29:13.

[108] Katopodis N, Hirshaut Y, Geller NL, Stock CC. Lipid-associated sialic acid test for the detection of human cancer. Cancer Research. 1982;42:5270-5.

[109] Plucinsky MC, Michael Riley W, Prorok JJ, Alhadeff JA. Total and lipid‐associated serum sialic acid levels in cancer patients with different primary sites and differing degrees of metastatic involvement. Cancer. 1986;58:2680-5.

[110] Cummings RD, Etzler M. Essentials of glycobiology: Cold Spring Harbor Laboratory Press; 2009.

[111] Duan X, Cai L, Lee LA, Chen H, Wang Q. Incorporation of azide sugar analogue decreases tumorigenic potential of breast cancer cells by reducing cancer stem cell population. Science China Chemistry. 2013;56:279-85.

[112] Vasiliou V, Nebert DW. Analysis and update of the human aldehyde dehydrogenase (ALDH) gene family. Hum Genomics. 2005;2:138-43.

[113] Marchitti SA, Brocker C, Stagos D, Vasiliou V. Non-P450 aldehyde oxidizing enzymes: the aldehyde dehydrogenase superfamily. Expert Opinion on Drug Metabolism and Toxicology. 2008;4:697-720.

[114] Marchitti SA, Brocker C, Stagos D, Vasiliou V. Non-P450 aldehyde oxidizing enzymes: the aldehyde dehydrogenase superfamily. 2008.

[115] Black WJ, Stagos D, Marchitti SA, Nebert DW, Tipton KF, Bairoch A, et al. Human aldehyde dehydrogenase genes: alternatively spliced transcriptional variants and their suggested nomenclature. Pharmacogenet Genomics. 2009;19:893-902.

[116] Muto M, Hitomi Y, Ohtsu A, Ebihara S, Yoshida S, Esumi H. Association of aldehyde dehydrogenase 2 gene polymorphism with multiple oesophageal dysplasia in head and neck cancer patients. Gut. 2000;47:256-61.

[117] Yokoyama A, Muramatsu T, Omori T, Yokoyama T, Matsushita S, Higuchi S, et al. Alcohol and aldehyde dehydrogenase gene polymorphisms and oropharyngolaryngeal, esophageal and stomach cancers in Japanese alcoholics. Carcinogenesis. 2001;22:433-9.

[118] Hanahan D, Weinberg RA. The hallmarks of cancer. Cell. 2000;100:57-70.

[119] Tlsty TD, Coussens LM. Tumor stroma and regulation of cancer development. Annual Review of Phytopathology. 2006;1:119-50.

[120] Reya T, Morrison SJ, Clarke MF, Weissman IL. Stem cells, cancer, and cancer stem cells. Nature. 2001;414:105-11.

[121] Ma I, Allan AL. The role of human aldehyde dehydrogenase in normal and cancer stem cells. Stem Cell Reviews and Reports. 2011;7:292-306.

[122] Park CH, Bergsagel DE, McCulloch EA. Mouse myeloma tumor stem cells: a primary cell culture assay. Journal of the National Cancer Institute. 1971;46:411-22.

[123] Bruce WR, Van Der Gaag H. A quantitative assay for the number of murine lymphoma cells capable of proliferation in vivo. Nature. 1963;199:79-80.

[124] Lapidot T, Sirard C, Vormoor J, Murdoch B, Hoang T, Caceres-Cortes J, et al. A cell initiating human acute myeloid leukaemia after transplantation into SCID mice. Nature. 1994;367:645-8.

[125] Al-Hajj M, Wicha MS, Benito-Hernandez A, Morrison SJ, Clarke MF. Prospective identification of tumorigenic breast cancer cells. Proceedings of the National Academy of Sciences USA. 2003;100:3983-8.

[126] Brennan SK, Meade B, Wang Q, Merchant AA, Kowalski J, Matsui W. Mantle cell lymphoma activation enhances bortezomib sensitivity. Blood. 2010;116:4185-91.

[127] Collins AT, Berry PA, Hyde C, Stower MJ, Maitland NJ. Prospective identification of tumorigenic prostate cancer stem cells. Cancer Research. 2005;65:10946-51.

[128] Li C, Heidt DG, Dalerba P, Burant CF, Zhang L, Adsay V, et al. Identification of pancreatic cancer stem cells. Cancer Research. 2007;67:1030-7.

[129] Lingala S, Cui YY, Chen X, Ruebner BH, Qian XF, Zern MA, et al. Immunohistochemical staining of cancer stem cell markers in hepatocellular carcinoma. EXxperimental and Molecular Pathology. 2010;89:27-35.

[130] Lugli A, Iezzi G, Hostettler I, Muraro MG, Mele V, Tornillo L, et al. Prognostic impact of the expression of putative cancer stem cell markers CD133, CD166, CD44s, EpCAM, and ALDH1 in colorectal cancer. British Journal of Cancer. 2010;103:382-90.

[131] O'Brien CA, Pollett A, Gallinger S, Dick JE. A human colon cancer cell capable of initiating tumour growth in immunodeficient mice. Nature. 2007;445:106-10.

[132] Ricci-Vitiani L, Lombardi DG, Pilozzi E, Biffoni M, Todaro M, Peschle C, et al. Identification and expansion of human colon-cancer-initiating cells. Nature. 2007;445:111-5.

[133] Yang ZF, Ho DW, Ng MN, Lau CK, Yu WC, Ngai P, et al. Significance of CD90+ cancer stem cells in human liver cancer. Cancer Cell. 2008;13:153-66.

[134] Ma I, Allan AL. The role of human aldehyde dehydrogenase in normal and cancer stem cells. Stem Cell Rev. 2011;7:292-306.

[135] Jones RJ, Barber JP, Vala MS, Collector MI, Kaufmann SH, Ludeman SM, et al. Assessment of aldehyde dehydrogenase in viable cells. Blood. 1995;85:2742-6.

[136] Storms RW, Trujillo AP, Springer JB, Shah L, Colvin OM, Ludeman SM, et al. Isolation of primitive human hematopoietic progenitors on the basis of aldehyde

dehydrogenase activity. Proceedings of the National Academy of Sciences USA. 1999;96:9118-23.

[137] Russo JE, Hauquitz D, Hilton J. Inhibition of mouse cytosolic aldehyde dehydrogenase by 4-(diethylamino) benzaldehyde. Biochemical Pharmacology. 1988;37:1639-42.

[138] Charafe-Jauffret E, Ginestier C, Iovino F, Wicinski J, Cervera N, Finetti P, et al. Breast cancer cell lines contain functional cancer stem cells with metastatic capacity and a distinct molecular signature. Cancer Research. 2009;69:1302-13.

[139] Mullendore ME, Koorstra JB, Li YM, Offerhaus GJ, Fan X, Henderson CM, et al. Ligand-dependent Notch signaling is involved in tumor initiation and tumor maintenance in pancreatic cancer. Clinical Cancer Research. 2009;15:2291-301.

[140] Huang EH, Hynes MJ, Zhang T, Ginestier C, Dontu G, Appelman H, et al. Aldehyde dehydrogenase 1 is a marker for normal and malignant human colonic stem cells (SC) and tracks SC overpopulation during colon tumorigenesis. Cancer research. 2009;69:3382-9.

[141] Corti S, Locatelli F, Papadimitriou D, Donadoni C, Salani S, Del Bo R, et al. Identification of a primitive brain-derived neural stem cell population based on aldehyde dehydrogenase activity. Stem Cells. 2006;24:975-85.

[142] Jiang F, Qiu Q, Khanna A, Todd NW, Deepak J, Xing L, et al. Aldehyde dehydrogenase 1 is a tumor stem cell-associated marker in lung cancer. Molecular Cancer Research. 2009;7:330-8.

[143] Werner Goedde H, Agarwal DP. Pharmacogenetics of aldehyde dehydrogenase (ALDH). Pharmacology & Therapeutics. 1990;45:345-71.

[144] Klyosov AA, Rashkovetsky LG, Tahir MK, Keung W-M. Possible role of liver cytosolic and mitochondrial aldehyde dehydrogenases in acetaldehyde metabolism. Biochemistry. 1996;35:4445-56.

[145] Li Y, Zhang D, Jin W, Shao C, Yan P, Xu C, et al. Mitochondrial aldehyde dehydrogenase-2 (ALDH2) Glu504Lys polymorphism contributes to the variation in efficacy of sublingual nitroglycerin. Journal of Clinical Investigation. 2006;116:506-11.

[146] Jo SA, Kim EK, Park MH, Han C, Park HY, Jang Y, et al. A Glu487Lys polymorphism in the gene for mitochondrial aldehyde dehydrogenase 2 is associated with myocardial infarction in elderly Korean men. Clinica Chimica Acta. 2007;382:43-7.

[147] Hui P, Nakayama T, Morita A, Sato N, Hishiki M, Saito K, et al. Common single nucleotide polymorphisms in Japanese patients with essential hypertension: aldehyde dehydrogenase 2 gene as a risk factor independent of alcohol consumption. Hypertension Research. 2007;30:585-92.

[148] Maring JA, Deitrich RA, Little R. Partial purification and properties of human brain aldehyde dehydrogenases. Journal of Neurochemistry. 1985;45:1903-10.

[149] Kamino K, Nagasaka K, Imagawa M, Yamamoto H, Yoneda H, Ueki A, et al. Deficiency in mitochondrial aldehyde dehydrogenase increases the risk for late-onset Alzheimer's disease in the Japanese population. Biochemical and Biophysical Research Communications. 2000;273:192-6.

[150] Wang B, Wang J, Zhou S, Tan S, He X, Yang Z, et al. The association of mitochondrial aldehyde dehydrogenase gene (ALDH2) polymorphism with susceptibility to late-onset Alzheimer's disease in Chinese. Journal of the Neurological Sciences. 2008;268:172-5.

[151] Matsuda T, Yabushita H, Kanaly RA, Shibutani S, Yokoyama A. Increased DNA damage in ALDH2-deficient alcoholics. Chem Res Toxicol. 2006;19:1374-8.

[152] Giorgianni F, Bridson PK, Sorrentino BP, Pohl J, Blakley RL. Inactivation of aldophosphamide by human aldehyde dehydrogenase isozyme 3. Biochemical Pharmacology. 2000;60:325-38.

[153] Magni M, Shammah S, Schiro R, Mellado W, Dalla-Favera R, Gianni AM. Induction of cyclophosphamide-resistance by aldehyde-dehydrogenase gene transfer. Blood. 1996;87:1097-103.

[154] Dipple KM, Crabb DW. The mitochondrial aldehyde dehydrogenase gene resides in an HTF island but is expressed in a tissue-specific manner. Biochemical and Biophysical Research Communications. 1993;193:420-7.

[155] Stewart MJ, Malek K, Crabb DW. Distribution of messenger RNAs for aldehyde dehydrogenase 1, aldehyde dehydrogenase 2, and aldehyde dehydrogenase 5 in human tissues. Journal of Investigative Medicine. 1996;44:42-6.

[156] Murata M, Tagawa M, Watanabe S, Kimura H, Takeshita T, Morimoto K. Genotype difference of aldehyde dehydrogenase 2 gene in alcohol drinkers influences the incidence of Japanese colorectal cancer patients. Japanese Journal of Cancer Research. 1999;90:711-9.

[157] Livak KJ, Schmittgen TD. Analysis of relative gene expression data using Real-Time Quantitative PCR and the $2^{-\Delta\Delta CT}$ method. Methods. 2001;25:402-8.

[158] Keung W-M, Vallee BL. Daidzin: a potent, selective inhibitor of human mitochondrial aldehyde dehydrogenase. Proceedings of the National Academy of Sciences. 1993;90:1247-51.

[159] Shirota F, DeMaster E, Kwon C, Nagasawa H. Metabolism of cyanamide to cyanide and an inhibitor of aldehyde dehydrogenase (ALDH) by rat liver microsomes. Alcohol and Alcoholism (Oxford, Oxfordshire). 1986;1:219-23.

[160] Pacini S, Carnicelli V, Trombi L, Montali M, Fazzi R, Lazzarini E, et al. Constitutive expression of pluripotency-associated genes in mesodermal progenitor cells (MPCs). PloS One. 2010;5:E9861.

[161] Pirozhkova I, Barat A, Dmitriev P, Kim E, Robert T, Guégan J, et al. Differences in transcription patterns between induced pluripotent stem cells produced from the same germ layer are erased upon differentiation. PloS One. 2013;8:E53033.

[162] Metcalfe C, de Sauvage FJ. A tumor-specific stem cell. Nature Genetics. 2013;45:7-9.

[163] Perry SW, Norman JP, Barbieri J, Brown EB, Gelbard HA. Mitochondrial membrane potential probes and the proton gradient: a practical usage guide. Biotechniques. 2011;50:98.