An RNA aptamer against a cancer stem cell marker epithelial cell adhesion molecule (EpCAM)

Sarah Shigdar[1], Jia Lin[1], Yan Yu[1], Mile Pastuovic[1], M Q Wei[2], and W Duan[*1]

[1]School of Medicine, Deakin University, Waurn Ponds, Victoria, 3217, Australia;

[2]Division of Molecular and Gene Therapies, Griffith Health Institute, School of Medical Science, Griffith University, Gold Coast campus, Southport, Queensland, 4215, Australia;

*Correspondence: Dr W Duan; E-mail: wduan@deakin.edu.au; School of Medicine, Deakin University, Pigdons Road, Warn Ponds, Victoria, 3217, Australia;

Running title: RNA Aptamer against EpCAM

The lack of a specific targeting strategy against cancer stem cells in current cancer treatment regimens is at least partly responsible for life-threatening cytotoxicity in patients undergoing traditional chemotherapy and the failure of the modern medicine in providing a cure for solid cancers. An effective cancer stem cell targeting system is urgently required for improved cancer therapy and molecular diagnosis. Epithelial cell adhesion molecule (EpCAM) is overexpressed in most solid cancers and it has recently been identified as a cancer stem cell marker. In this study, we isolated a 40 base RNA aptamer that binds to EpCAM from a random oligonucleotide library by exponential enrichment using systematic evolution of ligands. The aptamer was further truncated to 19 bases long. This EpCAM RNA aptamer interacts specifically with a number of live human cancer cells derived from breast, colorectal and gastric cancers that express EpCAM but not with those not expressing EpCAM, as analysed using flow cytometry and laser scanning confocal microscopy. The K_D of the EpCAM RNA aptamer to live human cancer cells is approximately 55 nM which, importantly, is efficiently internalised upon binding to cell surface EpCAM. To our knowledge, this is the first RNA aptamer against a cancer stem cell surface marker being reported. Such cancer stem cell aptamers will greatly facilitate the development of novel targeted nanomedicine and molecular imaging agents for cancer theranostics.

Introduction

The epithelial cell adhesion molecule, EpCAM (also known as CD326 or ESA), is a pleiotropic molecule, capable of both promoting and preventing epithelial cell-cell adhesion [1]. It is a 30-40 KDa type I glycosylated membrane protein expressed at a low level in a variety of human epithelial tissues. EpCAM is overexpressed in most solid cancers [2-4]. For example, intense expression of EpCAM is found in more than 98% patients with colorectal cancer [5]. Two decades of studies have shed light on the roles that EpCAM plays in tumourigenesis. Rather than antagonising apoptosis, EpCAM acts by inducing proliferation with a direct impact on cell cycle control, up-regulating the proto-oncogene c-myc and cyclins A and E, and signal transduction into the cell nucleus via the *wnt* pathway [2, 3, 6-8].

Recently, it has been recognised that a small proportion of cancer cells possess unlimited proliferation potential and are able to self-renew and to generate differentiated cancer cell progeny. These so-called cancer stem cells (also known as cancer initiating cells) are resistant to chemotherapy and radiotherapy [9]. It is thought that cytotoxic drugs and radiation kill mainly the bulk tumour cells but spare the cancer stem cells and thus a cure or even long-term control of macroscopic solid cancers by chemotherapy is still an exception rather than the rule. Therefore, in order to eradicate cancer, one must target and eliminate cancer stem cells.

EpCAM has been identified to be a cancer stem cell marker in a number of solid cancers, including breast cancer [10], colorectal cancer [4], pancreatic cancer [11] and liver cancer [12]. It has been shown that the expression of EpCAM in various cancers is inversely related to the prognosis of the patients [13]. Anti-EpCAM antibodies can be used to identify circulating tumour cells (metastatic cells) in the blood for cancers of the breast prostate, pancreas,

stomach and lung to provide prognostic information enabling individualized treatment of cancer [14]. However, initial clinical trials with anti-EpCAM antibodies failed to demonstrate objective clinical responses. [15-17]. It's thought that the large size of the antibody is a limitation to the distribution and delivery of monoclonal antibodies [15]. In addition, the antibody-dependent cytotoxicity relies on the carbohydrate composition in the CH2 domain of the antibody, which can vary significantly during antibody production. Thus, a smaller and more effective EpCAM targeting molecule is needed for targeted cancer therapy.

Aptamers, also known as chemical antibodies, are short single-stranded DNA or RNA. They fold into a three-dimensional structure and are capable of binding to target molecules with high affinity and selectivity. They are advantageous over antibodies in cancer targeting as they can be chemically synthesised with low batch variation and are cheaper to manufacture in large quantities. They are also more stable, nonimmunogenic and nontoxic. Because aptamers are 10 times smaller than an antibody, they exhibit superb tissue penetration properties. Produced via a method known as SELEX (systematic evolution of ligands by exponential enrichment) [18, 19], one aptamer has already been approved by the US Food and Drug Administration [20] for clinical use, and several aptamers are in clinical trials [21, 22].

Given the varying success of current anti-EpCAM antibody therapy, we sought to develop a nuclease resistant aptamer targeting the cancer stem cell marker that could have diagnostic and therapeutic potential. Herein, we describe the selection and characterisation of an RNA aptamer against EpCAM.

Materials and Methods

Cell lines and cell culture. The cell lines of human origin used in this study were purchased from American type Culture Collection. They are human gastric cancer Kato III; glioblastoma U118-MG; breast cancer MCF-7, MDA-MB-231 and T47D; colorectal cancer HT-29 and SW480; neuroblastoma SK-N-DZ, chronic myelogenous leukemia K562 and embryonic kidney cells HEK293T. Cells were grown and maintained in culture with Dulbecco's Modified Eagle medium (DMEM) (Invitrogen, Victoria, Australia) supplemented with 10 % fetal calf serum (HEK293T, MCF-7, SW480, T47D, HT-29, SK-N-DZ, K562, and U118-MG), or modified Iscove's minimal essential medium (IMEM) (Invitrogen) with 20% fetal calf serum (Kato III) or Roswell Park Memorial Institute 1640 (RPMI 1640) (Invitrogen) with 10% fetal calf serum (MDA-MB-231). Cells were maintained at 37 °C in a 5% CO_2 atmosphere.

Protein Expression and Preparation. Human EpCAM cDNA was purchased from Invitrogen and cloned into a mammalian expression vector, pcDNA 3.1/V5-His-TOPO. The recombinant 6xHistine-tagged EpCAM was transiently expressed in HEK293T and the total cell lysate was prepared as described previously [23]. Each well in the DELFIA anti-mouse-IgG coated plate (PerkinElmer, Cat. No. #4007-0010) was incubated with 1 µg of anti-His monoclonal antibody in binding buffer (Dulbecco's phosphate buffered saline (DPBS) containing 5 mM $MgCl_2$, 0.1 mg/mL tRNA and 0.1 mg/mL Salmon sperm) for 1 h and blocked with SuperBlock (Pierce) for 1 h at 23 °C followed by extensive binding in wash buffer. The cell lysate containing the recombinant EpCAM was added to the wells, incubated for 1 h at 37 °C followed by three washes. The EpCAM strips were kept moist at 4 °C until use.

SELEX selection. A DNA library containing a central 40-nt randomised sequence (5'-<u>TAA TAC GAC TCA CTA TAG</u> GGA CAC AAT GGA CG-N40-TAA CGG CCG ACA TGA GAG-3', with the T7 RNA polymerase promoter sequence underlined) was synthesised (GeneWorks, Australia). The double stranded DNA pool was generated from the original synthetic library via a large scale PCR using primers flanking the randomised sequence, 5'-TAA TAC GAC TCA CTA TAG GGA CAC AAT GGA CG-3' and 5'-CTC TCA TGT CGG CCG TTA-3'. A portion of the large-scale PCR products (~ 10^{14} sequences) was used as a template for in vitro transcription to produce the initial 2'-fluoropyrimidine modified RNA pool using a Durascribe® T7 Transcription kit (EPICENTRE® Biotechnologies, USA). For SELEX, RNA, at a concentration of 5 μM for initial selection or 1 μM for each iterative rounds, was diluted in 100 μL of binding buffer (Dulbecco's phosphate buffered saline containing 5 mM MgCl2, 0.1 mg/mL tRNA and 0.1 mg/mL Salmon sperm) and denatured at 85 °C for 5 minutes, allowed to cool to room temperature for 10 min, and annealed at 37 °C for 15 min, before incubating with the target protein in each well for 1 h at 23 °C. Following incubation and extensive washes, the bound RNA was reverse transcribed *in situ* using SuperScript III Reverse Transcriptase (Invitrogen), followed by PCR amplification and in vitro transcription and used for the next round of SELEX. Counter-selection steps were included from round 4, using a His-tagged irrelevant protein, to decrease the enrichment of species specifically recognising the His-tag, the antibodies or the well-coating. The number of PCR amplification cycles was also optimised to prevent over-amplification of non-specific "parasite" PCR products. In addition, the stringency of the selection process was enhanced to promote the selection of high-affinity aptamers through adjustments to aptamer concentration, incubation times, and the number of washes. To acquire aptamers of high specificity, the number of wells of protein targets used was progressively decreased while the washing stringency increased during the progression of SELEX, with negative selections

included from round four. Enrichment was monitored using restriction fragment length polymorphism (RFLP) and flow cytometry using live cells.

RFLP analysis. The enrichment of aptamer candidates during selection was determined by RFLP. Briefly, RFLP was performed as previously described [24, 25], with minor modifications. Approximately 5 ng of cDNA from iterative cycles was amplified by PCR for eight cycles. The amplified DNA was digested with 4 restriction enzymes, *Afa* I, *Alu* I, *Hha* I and *Xsp* I that recognise 4 nucleotides (frequent cutters) in Buffer T supplied by the manufacturer (Takara) with 0.1% (w/v) bovine serum albumin at 37 °C overnight. Following the overnight digestion, the DNA was heated to 65 °C, cooled on ice, and separated via electrophoresis on a native 20% polyacrylamide gel in TBE buffer. The gel was then stained in GelStar and visualized using a standard gel imaging system.

Flow Cytometry assays. Cells were harvested at 80% confluence with trypsin digestion and resuspended in washing buffer (DPBS with 5 mM $MgCl_2$) and enumerated. Following centrifugation (1000 x g for 5 min), the pellet was resuspended in binding buffer and diluted to 1×10^6/mL.

To confirm aptamer binding to the target protein, RNA from iterative rounds were labelled at the 3'-ends with fluorescein isothiocyanate (FITC) according to a previously described method [26]. Amber tubes were used throughout to minimise photo-bleaching. Briefly, samples were oxidized with sodium periodate. The oxidation was terminated with the addition of 10 mM ethylene glycol, followed by ethanol precipitation. FITC was added at a 30-molar excess, and the reaction was completed overnight at 4°C. One µM of FITC-labelled RNA was incubated with trypsinised 5×10^5 Kato III or U118-MG cells in 100 µL of binding buffer for 1 h on ice, followed by washing three times and resuspension in 300 µL of binding buffer. Fluorescent intensity was determined with a FACS Canto II flow cytometer (Becton

Dickinson) by counting 50,000 events each sample. The FITC-labelled RNA from the unselected library and an EpCAM-negative cell line were used to determine non-specific binding.

The binding for each round was calculated after subtracting the mean fluorescence intensity of the binding of round zero RNA to target cells as well as that for binding to negative control cells according to a method described by Ellington and colleagues [27].

Cloning, Sequencing and Structural Analysis of Selected Aptamers. Following RFLP and flow cytometric analyses of iterative rounds, we found that round seven demonstrated sufficient enrichment of RNA sequences that selectively recognised the target protein. This enriched pool was amplified by PCR for ten cycles and the PCR products were cloned into the plasmid pCR®4-TOPO® (Invitrogen). Plasmid DNA from individual clones was prepared and its sequence determined using an automated DNA sequencing procedure. The aptamer sequences were analysed using ClustalX2 [28]. Secondary structures were predicted using the program RNAfold [29].

Determination of aptamer affinity. The dissociation constant (K_D) of successful 2'-fluoropyrimidine RNA aptamer species to native EpCAM expressed on the cell surface was determined using flow cytometry. Kato III or U118MG cells (5×10^5) were first incubated with blocking buffer (binding buffer containing 0.2 % (w/v) sodium azide) followed by two washes with binding buffer prior to incubation with serial concentrations (approximately 10-fold above and below the apparent K_D) of FITC-labelled aptamer in a 100 µL volume of binding buffer for 1 h on ice. The cells were washed three times with binding buffer, resuspended in 150 µL binding buffer and subjected to flow cytometric analyses. The FITC-labelled unselected library was used as another negative control. The mean fluorescence intensity (MFI) of the unselected library was subtracted from that of the aptamer-target cell to

generate the MFI of specific binding. The K_D for each aptamer was determined by Scatchard analysis according to the equation:

[Bound aptamer]/[aptamer]= -(1/K_D) × [bound aptamer] + ([T]$_{tot}$/K_D)

where [T]$_{tot}$ represents the total target concentration.

Aptamer Truncation and Determination of Specificity. To generate the truncated aptamer, the sense and antisense DNA oligonucleotides of desired sequence were synthesised. EpDT1 (1st Truncation) derived from a sense oligonucleotide, 5'- <u>TAATACGACTCACTATA</u>GGTCCGTAGTTCTGGCTGACTGGTTACCCGGTCGTACA GCTCG-3', and antisense oligonucleotide, 5'-CGAGCTGTACGACCGGGTAACCAGTC AGCCAGAACTACGGACCTATAGTGAGTCGTATTA-3'; and EpDT3 (the third truncation) was derived from a sense oligonucleotide: 5'- <u>TAATACGACTCACTATA</u>GCGACTGGTT ACCCGGTCG-3' and an antisense oligonucleotide, 5'-CGACCGGGTAACCAGTCGCTA TAGTGAGTCGTATTA-3' (T7 RNA promoter sequence is underlined). The pair of oligonucleotides was mixed in equal molar ratios in 1 × PEI buffer (0.1 M Tris-HCl pH 8, 0.1 M MgCl$_2$ 0.5 M NaCl and 0.1 M dithiothreitol), heated for 5 min at 90°C and cooled slowly to room temperature prior to ethanol precipitation. In vitro transcription and FITC-labelling was performed as described above. The final truncation of this clone (EpDT3), 5'-DY647- GCGACUGGUUCCCGGUCGdT-3', was also chemically synthesized with a 5'-DY647 fluorescent tag and a 3'-inverted deoxythymidine (Dharmacon). The binding affinity of DY647-EpDT3 was determined as described above, using EpCAM-positive and -negative cell lines [30-32] and a negative control aptamer

5'-DY647-mGCmGACUmGmGUUmACCCmGmGUCmGdT-3'. The blocking step was performed at 4 °C using blocking buffer containing 5 % (v/v) fetal calf serum, whilst the binding of the aptamers was performed at 37 °C for 30 min.

Confocal microscopy. Twenty-four hours prior to labelling, cells were seeded at a density of 75,000 cells per cm^2 in an 8-chamber slide (Lab-Tek I, Nunc). DY647-EpDT3 and the control aptamer were prepared in the same manner as for flow cytometry. Following removal of media, cells were incubated in blocking buffer containing 5 % (v/v) serum at 37 °C for 15 min, washed twice in binding buffer containing 0.2 % (w/v) sodium azide prior to incubation with 100 nM aptamer for 30 min at 37 °C. Bisbenzimide Hoechst 33342 (3 µg/mL) (Sigma) was added to the cells during the final 15 mins of incubation. The aptamer solution was removed and the cells washed 5 min each in binding buffer for 3 times prior to visualisation using a FluoView FV10i laser scanning confocal microscope (Olympus).

Inhibition of endocytosis. This was performed essentially as described for confocal microscopy with minor modifications. Briefly, cells were pre-treated with either a potassium-depleted (50 mM Hepes, 140 mM NaCl, 2.5 mM $MgCl_2$, and 1 mM $CaCl_2$) or a hypertonic buffer (potassium-depleted buffer containing 3 mM KCl and 450 mM sucrose) for 1 hr at 37° C. These buffers were also used in the incubation step with aptamers and all rinsing steps. The effectiveness of these treatments in inhibiting endocytosis was evaluated by qualitatively characterising the internalisation of human transferrin conjugated to Alexa Fluor 488 (Invitrogen). Transferrin (5 µg/mL) was added to the cells following pre-treatment followed by a 30 min incubation at 37° C. The cells were washed three times in their respective buffers and visualised using the FluoView FV10i confocal microscope.

Western Analysis. The expression of EpCAM protein in cultured cells was analysed by Western analysis as previously described [23]. Fifteen microlitres of each sample was then

loaded onto a 12% NuPAGE Bis-Tris mini gel (Invitrogen) along with a Precision Plus dual colour protein standard (BioRad). Following electrophoresis for 45 min at 200V, the protein was transferred to a nitrocellulose membrane (Invitrogen) and blocked with 5% skimmed milk for 3 h at 25°C, before being incubated with either anti-β-actin (Sigma) diluted 1:2000, or the anti-EpCAM antibodies, AUA1 (Abcam) diluted 1:500, or 323/A3 (Abcam) diluted 1:250 in 1% skimmed milk, overnight at 4°C. Chemiluminescence was detected using an ImageQuant™ LAS 4000 Biomolecular Imager (GE Healthcare).

Results

A robust SELEX and RFLP facilitate selection of EpCAM aptamers

To select RNA aptamers that recognise cell surface EpCAM of native confirmation, we devised a modified SELEX procedure as described in Materials and Methods. In our SELEX procedure, the cytoplasmic domain (the C-terminal of EpCAM) was attached to the solid support leaving the extracellular domain of the transmembrane protein (the N-terminal of EpCAM) freely exposed to the selection solution to facilitate the selection of aptamers that bind to the extracellular domain of the target protein. A random RNA library of approximately 10^{15} species containing 2'-fluoro-modified ribose on all pyrimidines was used to incubate with immobilised His-tagged recombinant EpCAM. After washing to remove unbound RNA, the bound RNA was amplified via reverse transcription coupled to PCR (RT-PCR) in situ. In total, 12 iterative rounds of SELEX were performed and enrichment was monitored through the use of both non-radioactive restriction fragment length polymorphism (RFLP) and flow cytometry of live cells. RFLP permits quick evaluation of our SELEX rounds and allowed us to shorten the number of SELEX cycles required as enrichment was observed from cycle six. Enzyme digestion of PCR products from iterative rounds showed an evolving pattern of digestion (Fig. 1). The digestion products from the first few rounds of SELEX showed a heterogeneous composition, with none or scarce digestion observed. Digestion appeared to peak at rounds six to eight, before losing some integrity in the 75 bp product from round nine. The digestion pattern continued to devolve, with the smallest digestion product being lost by round twelve. In addition to RFLP, the ability of selected RNA aptamers to interact specifically to EpCAM-positive human cancer cells but not to other sites on a human cell surface non-specifically was assessed using live cells by flow cytometry. Fig. 2 shows that RNA aptamers from SELEX round seven bound 2.2-fold higher

to live EpCAM-positive human gastric cancer cells compared with that of EpCAM-negative live human glioblastoma cells and to our unselected library.

Generation of a small EpCAM aptamer via post-SELEX engineering. Following successful enrichment of aptamers to EpCAM, individual aptamer was cloned and sequenced. A total of 60 clones were sequenced from rounds seven and eleven, and sequence homology was determined using ClustalX2. The binding affinity of six distinct families was analysed using fluorescein isothiocyanate (FITC)-labelled RNA on EpCAM-positive human cancer cells and EpCAM–negative cell human cell lines (data not shown). The EpCAM aptamer with the highest affinity, Clone D (GGG ACA CAA UGG ACG UCC GUA GUU CUG GCU GAC UGG UUA CCC GGU CGU ACA GCU CUA ACG GCC GAC AUG AGA G), is 73-nucleotide long and it was found to have a K_D of 278 nM (Fig. 3a) in the binding assay using human gastric carcinoma cells Kato III. For in vivo cancer molecular imaging and targeted therapy, an aptamer of small size is desirable. Thus, we performed serial truncation of the original clone D EpCAM aptamer, based on the assumption that the end loop on the left of the 2-D structure of Clone D (Fig 3A, inset) was responsible for the binding of the aptamer to EpCAM. Two rounds of truncation shortened the Clone D first to 43 bases and then further down to 19 bases (Fig 3b, & c, inset). The equilibrium dissociation constants (K_D) for the shortest FITC-labelled EpCAM aptamer, EpDT3 (5'-GCGACUGGUUACCCGGUCG-3), was 18 nM, when analysed using Kato III cells (Fig 3c).

The size of the fluorescent tag influences binding affinity. All the RNA aptamers described in the proceeding sections were produced by enzymatic in vitro transcription, which contains a section of T7 RNA polymerase promoter sequence at the 5'-end of the RNA. In order to study the utility of our aptamers in targeted cancer nanomedicine in vivo, we switched to the approach of total chemical synthesis of RNA aptamers. We synthesised our RNA aptamer with a 3'-inverted deoxythymidine cap, leading to a 3'-3' linkage which

inhibits degradation by 3' exonucleases. This aptamer was also labelled with a different fluorescent tag on the 5' end of the oligonucleotide. The DY647 was chosen for future downstream applications, having a far-red emission spectra, allowing us to identify its biodistribution profile in animal studies. Interestingly, the apparent K_D was increased, from 12.0 nM to 54.5 nM with this Dy647-labelled EpDT3. This could due to several factors, including the different size of the fluorophore (389 Da for FITC vs 1008 Da for Dy647) and the presence of an inverted dT or FITC at the 3'-end rather than a DY647 at the 5' end.

Ep-DT3-DY647 binds to EpCAM-positive human cancer cells highly specifically. Kato III cells, previously shown by others to possess a high number of EpCAM molecules on its cellular surface, were initially used to determine affinity of our aptamer to EpCAM. To test whether the EpCAM aptamer is able to bind EpCAM on the surface of human cancer cells of different histopathological origins, we studied the interaction of DY647-EpDT3 with a panel of EpCAM-positive human cancer cell lines using flow cytometry. Apart from Kato III, we used a breast adenocarcinoma-derived cell line MCF-7, a colon adenocarcinoma-derived cell line SW480, a ductal breast epithelial tumour-derived cell line T47D, a colon adenocarcinoma grade II-derived cell line HT-29, and a breast adenocarcinoma-derived cell line MDA-MB-231. As shown in Fig. 4a-4f, DY647-EpDT3 was able to bind to 6 different types of human cancer cells well, with an apparent K_D of approximately 60 nM. To confirm the observed interaction is specific between EpDT3 and the testing cancer cells, we performed the binding assay using a control DY647-labelled aptamer that has the same nucleotide sequence as EpDT3 but with a 2'-methyl (2'-Me) modification instead of a 2'-fluro (2'-F) in all the pyrimidines. As demonstrated in Fig. 4g-4i, the DY647-control aptamer failed to bind the six EpCAM-positive cancer cell lines. In order for this novel EpCAM aptamer to become an effective cancer targeting agent, it must have negligible interaction with cells that do not express EpCAM. To this end, we used a number of

different types of human cells, both non-transformed or cancerous, to study the specificity and selectivity of the EpCAM aptamer. As illustrated in Fig. 4m-4o, DY647- EpDT3 does not bind to EpCAM-negative human cells, such as a human embryonic kidney 293T cell line HEK293T, a neuroblastoma-derived cell line SK-N-DZ or an erythroleukaemia-derived cell line K562.   The specific interaction between EpDT3 RNA aptamer and EpCAM-positive cells but not with EpCAM-negative cell lines were further verified using flow cytometry analysis (Fig 4p, 4q) and confocal microscopy (Fig. 5a and 5b).

EpDT3-DY647 is internalised via receptor-mediated endocytosis. For cancer therapy and molecular imaging, the targeting moiety and/or its conjugated nanoparticle should ideally be internalised upon binding of the cancer cells. We investigated the fate of our aptamer after binding with cancer cells using confocal microscopy. We asked whether EpDT3-DY647 stays on the cell surface, or was internalised upon binding. We shoed that after binding, the pattern of red fluorescence (for EpDT3-DY647) displayed a particular intracellular pattern (Fig. 5a, the two upper panels), indicative of an entry via endocytosis. Interestingly, this cell internalisation is specific for EpDT3-DY647 as the control 2'-Me aptamer did not bind to these 6 cancer cell lines (Fig. 5, lower panels). To confirm this, we pre-treated cells with agents known to block endocytosis, i.e. potassium-depletion and hypertonic treatment, [33-35]. The effectiveness of such treatment in blocking endocytosis was first confirmed in these cells to block the internalisation of transferrin (Supp Fig 1).   Indeed, upon treatment with endocytosis blockers, a ring pattern of the red fluorescence for EpDT3-DY647 was observed in the four different human cancer cell lines tested, including human gastric, colorectal and breast cancers (Fig. 5c).   Thus, upon binding cell surface EpCAM, our aptamer was internalised by human cancer cells, possibly via clathrin-dependent endocytosis.

Discussion

Novel approaches for early detection and effective treatment of cancer will significantly improve clinical outcomes of cancer treatment. EpCAM is an attractive target for novel cancer therapy because it is overexpressed in most human adenocarcinomas and it is a marker for cancer stem cells in at least several solid cancers [2, 5]. However, the results from clinical trials with at least five different monoclonal antibodies against EpCAM as tumour immunotherapies were disappointing [21, 36-38]. The lack of success of EpCAM antibody monotherapy has been attributed, at least in part, to the size of the immunoglobulin, the affinity of the antibody, and its high immunogenicity [15, 17, 39]. Therefore, in this study, we attempted to isolate small RNA aptamers against EpCAM.

Through the use of RFLP and flow cytometric binding assays, we demonstrated the robustness of our aptamer selection regime which was successfully evolved from our randomised library and allowed us to clone the aptamers at seventh cycle of SELEX. Following evaluation of individual clones from round seven, and the establishment of a K_D of 211 nM for one clone, we further engineered this clone in an attempt to minimize the size of the aptamer. The resultant clone, EpDT3, had a K_D of approximately 12 nM. The K_D of all three versions of our aptamer are in good agreement with previously reported aptamers against other targets, such as the RNA aptamers to Tenascin-C (5 nM), *Trypanosoma cruzi* (172 nM), and Prostate-specific membrane antigen aptamer (2-11 nM) [40].

One of the interesting features of our aptamer is that it is able to bind to human cancer cells expression high levels as well as medium to low level of EpCAM. For example, DY647-EpDT3 bound to both gastric carcinoma cell Kato II which has approximately 893,100 EpCAM protein molecules per cell [30] and breast carcinoma cells MDA-MB-231 that has 1700 EpCAM protein molecules per cell [30] (Fig. 5). Our aptamer was shown to have a slightly different K_D towards different human cancer cell lines when analysed under physiological temperature and the concentration of magnesium (2.5 mM) (Fig. 4). These

results appear to be consistent with previous reports of aptamer binding to different cell lines. Indeed, one aptamer, KH1C12, had a K_D of 4.5 nM toward the HL60 cell line though a much lower recognition pattern for myeloid cell lines tested approximately 50% for Monomac-6, and MV4-11, and approximately 25% for Kasumi-1as compared to HL60) [41]. (Sarah to dig out exactly the Kd for MV4-11 if available). In a separate study for aptamers against glioma cells, a 2.5- to 6-fold difference in K_D for an aptamer GL56 was reported between glioma cell lines U87MG and SB-G [42].

One of the main advantages of aptamers over antibodies is that the former can be chemically synthesized without the need for using hybridoma or animals. While FITC (attached to the 3-end of the aptamer) proved a useful reporter molecule for initial binding assays in our characterization of EpCAM aptamers, we were interested in obtaining an aptamer that could be useful in molecular imaging. We therefore labelled the 5'-end of our chemically synthesized aptamer with a DY647 dye in the far-red spectrum. However, we did experience a four-fold decrease in binding affinity of DY647-EpDT3, as compared to EpDT3-FITC (54.5 nM versus 12 nM) at physiological temperature (Fig 4a vs Fig 3c). This could be due to several factors, the steric hindrance (molecular weight of 389 for FITC vs 1008 for Dy647), the position of the dye (3'- vs 5'-end) or presence or absence of the inverted dT cap. However, for clinical applications, the high affinity of a tumour targeting ligand may not necessarily be advantageous. In addition to being expressed in high levels in most human adenocarcinomas, EpCAM is also expressed at low levels in a number of normal epithelial cells, including gastrointestinal tract, bile ducts and pancreas [43]. Therefore, a high affinity EpCAM ligand may bind to EpCAM-positive tumours as well as normal epithelial cells. Indeed, recent clinical trials demonstrated that high-affinity humanized monocloncal antibodies to EpCAM, 362W94 (K_D = 0.19 nM) and ING-1 (K_D = 0.16 nM) caused acute pancreatitis in patients [44-46] ; while an antibody to EpCAM with moderate high affinity,

Adecatumumab (K_D = 91 nM), did not display such dose-limiting toxicity and was well tolerated in patients [39]. Importantly, our EpCAM aptamers do not bind with human cells that do not express EpCAM (Fig. Fig. 4m-4q, Fig. 5C). Thus, the DY647-EpDT3 with a moderate affinity could serve as a selective tumour target agent for cancer nanomedicine.

Having determined specificity and selectivity of our EpCAM aptamer, we sought to determine if this molecule is internalised upon binding its cell surface target. It is critically important for a cancer target ligand to be actively internalised into the cancer cells as this will enable the delivery and release of the payload, e.g. chemotherapy drugs, toxins and/or therapeutic radioisotopes, inside cancer cells, thus overcoming multidrug resistance and minimising the collateral killing of normal cells. The intracellular entrapment of a molecular imaging ligand will afford higher tumour to blood or tumour to normal tissue ratio. One of the disadvantages associated with aptamers, and nucleic acids in general, is that these molecules cannot be directly taken up into the cell due to the repulsion between the negatively charged aptamer and the cell membrane. In general, nucleic acid aptamers are not capable of being specifically internalised inside living cells, with a few exceptions being a DNA anti-PTK7 aptamer and RNA aptamers against PSMA, CD4 and HIV gp120 [47-49]. The final truncation of our aptamer, EpDT3 is 19 bases long and with a molecular mass of approximate 5.9 KDa. In this study, we showed that our anti-EpCAM aptamer is efficiently internalised following binding to the cell surface EpCAM (Fig. 5). We obtained evidence suggesting that our aptamer is internalised through an energy-dependent manner, as incubating cells at 4 °C or pretreating cells with 0.04% sodium azide significantly attenuated aptamer internalization (data not shown). We also subjected cancer cell lines to pretreatment of hypertonic sucrose and potassium depletion before incubation with the EpCAM aptamer (Fig. 5d). Hypertonic shock inhibits several pathways in addition to the clathrin-mediated endocytosis, while potassium depletion is a more selective inhibition to the lipid

raft/caveolae-mediated endocytosis [34]. Both of these treatments blocked the internalization of transferrin (Suppl Fig. 1) which is known to be internalized through clathrin-mediated endocytosis [33]. Therefore, it is plausible that upon binding cell surface EpCAM, the EpCAM aptamer is actively internalised via receptor-medicated encocytosis.

To conclude, we have developed the first RNA aptamer against a cancer stem cell marker that specifically binds to the cell surface EpCAM followed by active internalisation. This small and chemically synthesized cancer stem cell ligand will facilitate the development of novel targeted cancer nanomedicine and molecular imaging agents.

Figure Legends

Fig. 1. RFLP analysis of PCR products from the analysis of random RNA aptamer pools across the progressive rounds of SELEX. PCR products from each round of SELEX were digested with *Afa* I, *Alu* I, *Hha* I and *Xsp* I, followed by separation through 20% native polyacrylamide gel and stained GelStar®. SELEX round numbers are indicated above each lane, while the size of the molecular weight is indicated on the right.

Fig. 2. Isolation of EpCAM aptamers via SELEX. (**a**) Flow cytometric binding analysis of FITC-labelled aptamers from iterative rounds of SELEX to EpCAM-positive Kato III cells. Fluorescein-labelled RNA from each round was incubated with target cells followed by flow cytometric analysis. The binding for each round was calculated after subtracting the mean fluorescence intensity of the binding of unselected library RNA to target cells as well as that for binding to negative control cells. (**b**) Graphical representation of fold change of binding to EpCAM-positive Kato III cells versus EpCAM-negative U118-MG cells.

Fig. 3. Determination of equilibrium dissociation constants (K_D) for the interaction between serially truncated clones of EpCAM aptamers and Kato III cells. A representative binding curve at varying concentrations of EpCAM aptamers (1-1000 nM) at a Kato cell density of 5 x 10^5 cells/ml is shown. (**a**) Full-length EpCAM aptamer. (**b**) Aptamer from the first truncation, EpDT1. (c) Aptamer from the third truncation, EpDT3. The secondary structure

of each of the aptamers was modelled using RNAfold [29] and is shown under the respective binding curve.   (Sarah, we need to use the sequence including T7 pomoter sequence)

Fig. 4. Specificity of the EpCAM aptamer. DY-647-labelled EpDT3 was incubated with indicated human cell lines and analysed by flow cytometry. The mean fluorescence intensity (MFI) was plotted against varying concentrations of EpCAM aptamers (1-200 nM) at a cell density of 5 x 10^5 cells/ml. (**a-f**) Binding of aptamer to EpCAM-positive human cancer cells: Kato III (**a**), MCF-7 (**b**), SW480 (**c**), T47D (**d**), HT-29 (**e**) and MDA-MB-231 (**f**). (**g – l**) Binding of a DY-647-labelled negative control aptamer to EpCAM-positive cell lines: Kato III (**g**), MCF-7 (**h**), SW480 (**i**), T47D (**j**), HT-29 (**k**), and MDA-MB-231 (**l**). (**m-o**) Binding of DY647-labelled EpDT3 EpCAM-negative cell lines: HEK-293T (**m**), SK-N-DZ (**n**) and K562 (**o**). (**p – q**) Flow cytometric analysis of the binding of DY-647-labelled EpDT3 to EpCAM-positive cell lines versus that to the EpCAM-negative cell line, HEK-293T.

Fig 5.   The EpCAM RNA aptamer is endocytosed upon binding of cell surface EpCAM. DY-647-labelled EpCAM aptamer or control aptamer were incubated with indicated human cancer cells for 30 min at 37 °C followed by imaging using laser scanning confocal microscopy. For each pair of panels, optical (phase contrast) images are on the top while fluorescent images at the bottom. (**a**) Binding and subcellular distribution of aptamers to EpCAM-positive cancer cell lines. (**b**) Enlarged micrograph illustrating punctate pattern of fluorescently labelled EpCAM aptamer in a single Kato III cell. (**c**) Binding analysis of EpCAM aptamer to EpCAM-negative human cancer cell lines. (**d**) Binding and subcellular distribution of aptamers in EpCAM-positive cancer cell lines that had undergone potassium depletion or hypertonic treatment. Scale bars: 20 µM.

References

1. Bombardelli L, Cavallaro U. Immunoglobulin-like cell adhesion molecules: novel signaling players in epithelial ovarian cancer. Int J Biochem Cell Biol2010 May;42(5):590-4.

2. Munz M, Baeuerle PA, Gires O. The emerging role of EpCAM in cancer and stem cell signaling. Cancer research2009 Jul 15;69(14):5627-9.

3. Stoecklein NH, Siegmund A, Scheunemann P, *et al.* Ep-CAM expression in squamous cell carcinoma of the esophagus: a potential therapeutic target and prognostic marker. BMC Cancer2006;6:165.

4. Dalerba P, Dylla SJ, Park IK, *et al.* Phenotypic characterization of human colorectal cancer stem cells. Proceedings of the National Academy of Sciences of the United States of America2007 Jun 12;104(24):10158-63.

5. Went P, Vasei M, Bubendorf L, *et al.* Frequent high-level expression of the immunotherapeutic target Ep-CAM in colon, stomach, prostate and lung cancers. Br J Cancer2006 Jan 16;94(1):128-35.

6. Trzpis M, McLaughlin PM, de Leij LM, *et al.* Epithelial cell adhesion molecule: more than a carcinoma marker and adhesion molecule. Am J Pathol2007 Aug;171(2):386-95.

7. Maaser K, Borlak J. A genome-wide expression analysis identifies a network of EpCAM-induced cell cycle regulators. Br J Cancer2008 Nov 18;99(10):1635-43.

8. Maetzel D, Denzel S, Mack B, *et al.* Nuclear signalling by tumour-associated antigen EpCAM. Nat Cell Biol2009 Feb;11(2):162-71.

9. Hermann PC, Bhaskar S, Cioffi M, *et al.* Cancer stem cells in solid tumors. Semin Cancer Biol2010 Apr;20(2):77-84.

10. Al-Hajj M, Wicha MS, Benito-Hernandez A, *et al.* Prospective identification of tumorigenic breast cancer cells. Proceedings of the National Academy of Sciences of the United States of America2003 Apr 1;100(7):3983-8.

11. Li C, Heidt DG, Dalerba P, *et al.* Identification of pancreatic cancer stem cells. Cancer research2007 Feb 1;67(3):1030-7.

12. Yamashita T, Ji J, Budhu A, *et al.* EpCAM-positive hepatocellular carcinoma cells are tumor-initiating cells with stem/progenitor cell features. Gastroenterology2009 Mar;136(3):1012-24.

13. Baeuerle PA, Gires O. EpCAM (CD326) finding its role in cancer. Br J Cancer2007 Feb 12;96(3):417-23.

14. Mostert B, Sleijfer S, Foekens JA, *et al.* Circulating tumor cells (CTCs): detection methods and their clinical relevance in breast cancer. Cancer Treat Rev2009 Aug;35(5):463-74.

15. Schwartzberg LS. Clinical experience with edrecolomab: a monoclonal antibody therapy for colorectal carcinoma. Crit Rev Oncol Hematol2001 Oct;40(1):17-24.

16. Armstrong A, Eck SL. EpCAM: A new therapeutic target for an old cancer antigen. Cancer Biol Ther2003 Jul-Aug;2(4):320-6.

17. Oberneder R, Weckermann D, Ebner B, *et al.* A phase I study with adecatumumab, a human antibody directed against epithelial cell adhesion molecule, in hormone refractory prostate cancer patients. Eur J Cancer2006 Oct;42(15):2530-8.

18. Ellington AD, Szostak JW. In vitro selection of RNA molecules that bind specific ligands. Nature1990 Aug 30;346(6287):818-22.

19. Tuerk C, Gold L. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. Science1990 Aug 3;249(4968):505-10.

20. Lee JH, Canny MD, De Erkenez A, *et al.* A therapeutic aptamer inhibits angiogenesis by specifically targeting the heparin binding domain of VEGF165. Proceedings of the National Academy of Sciences of the United States of America2005 Dec 27;102(52):18902-7.

21. Das M, Mohanty C, Sahoo SK. Ligand-based targeted therapy for cancer tissue. Expert opinion on drug delivery2009 Mar;6(3):285-304.

22. Missailidis S, Hardy A. Aptamers as inhibitors of target proteins. Expert Opin Ther Pat2009 Aug;19(8):1073-82.

23. Yeong SS, Zhu Y, Smith D, *et al.* The last 10 amino acid residues beyond the hydrophobic motif are critical for the catalytic competence and function of protein kinase Calpha. J Biol Chem2006 Oct 13;281(41):30768-81.

24. Pestourie C, Cerchia L, Gombert K, *et al.* Comparison of different strategies to select aptamers against a transmembrane protein target. Oligonucleotides. [Article]. 2006;16(4):323-35.

25. Chauveau F, Aissouni Y, Hamm J, *et al.* Binding of an aptamer to the N-terminal fragment of VCAM-1. Bioorg Med Chem Lett2007 Nov 15;17(22):6119-22.

26. Willkomm D, Hartmann R. Handbook of RNA Biochemistry. In: Hartmann RK BA, Schön A, Westhof E, , editor. Weinheim: Wiley-VCH GmbH & Co. KGaA.; 2005. p. 86-94.

27. Li N, Ebright JN, Stovall GM, *et al.* Technical and biological issues relevant to cell typing with aptamers. J Proteome Res2009 May;8(5):2438-48.

28. Larkin MA, Blackshields G, Brown NP, *et al.* Clustal W and Clustal X version 2.0. Bioinformatics2007 Nov 1;23(21):2947-8.

29. Hofacker IL. Vienna RNA secondary structure server. Nucleic Acids Res2003 Jul 1;31(13):3429-31.

30. Prang N, Preithner S, Brischwein K, *et al.* Cellular and complement-dependent cytotoxicity of Ep-CAM-specific monoclonal antibody MT201 against breast cancer cell lines. Br J Cancer2005 Jan 31;92(2):342-9.

31. Osta WA, Chen Y, Mikhitarian K, *et al.* EpCAM is overexpressed in breast cancer and is a potential target for breast cancer gene therapy. Cancer research2004 Aug 15;64(16):5818-24.

32. Yu G, Zhang X, Wang H, *et al.* CpG island methylation status in the EpCAM promoter region and gene expression. Oncol Rep2008 Nov;20(5):1061-7.

33. Vercauteren D, Vandenbroucke RE, Jones AT, *et al.* The use of inhibitors to study endocytic pathways of gene carriers: optimization and pitfalls. Mol Ther2010 Mar;18(3):561-9.

34. Ivanov AI. Pharmacological inhibition of endocytic pathways: is it specific enough to be useful? Methods Mol Biol2008;440:15-33.

35. Altankov G, Grinnell F. Depletion of intracellular potassium disrupts coated pits and reversibly inhibits cell polarization during fibroblast spreading. J Cell Biol1993 Mar;120(6):1449-59.

36. McLaughlin PM, Kroesen BJ, Harmsen MC, *et al.* Cancer immunotherapy: insights from transgenic animal models. Crit Rev Oncol Hematol2001 Oct;40(1):53-76.

37. Wang X, Yang L, Chen ZG, *et al.* Application of nanotechnology in cancer therapy and imaging. CA: a cancer journal for clinicians2008 Mar-Apr;58(2):97-110.

38. White RR, Sullenger BA, Rusconi CP. Developing aptamers into therapeutics. Journal of Clinical Investigation. [Article]. 2000 Oct;106(8):929-34.

39. Schmidt M, Scheulen ME, Dittrich C, *et al.* An open-label, randomized phase II study of adecatumumab, a fully human anti-EpCAM antibody, as monotherapy in patients with metastatic breast cancer. Ann Oncol2010 Feb;21(2):275-82.

40. Wochner A, Menger M, Rimmele M. Characterisation of aptamers for therapeutic studies. Expert Opin Drug Discov2007;2(9):1205-24.

41. Sefah K, Tang ZW, Shangguan DH, *et al.* Molecular recognition of acute myeloid leukemia using aptamers. Leukemia. [Article]. 2009 Feb;23(2):235-44.

42. Cerchia L, Esposito CL, Jacobs AH, *et al.* Differential SELEX in human glioma cell lines. PLoS One2009;4(11):e7971.

43. Balzar M, Winter MJ, de Boer CJ, *et al.* The biology of the 17-1A antigen (Ep-CAM). J Mol Med1999 Oct;77(10):699-712.

44. Saleh MN, Posey JA, Khazaeli LM, *et al.* Phase I trial testing multiple doses of humanized monoclonal antibody (Mab) 3622W94. Proc Am Soc Clin Oncol1998;17:436.

45. Munz M, Murr A, Kvesic M, *et al.* Side-by-side analysis of five clinically tested anti-EpCAM monoclonal antibodies. Cancer Cell Int2010;10:44.

46. de Bono JS, Tolcher AW, Forero A, *et al.* ING-1, a monoclonal antibody targeting Ep-CAM in patients with advanced adenocarcinomas. Clin Cancer Res2004 Nov 15;10(22):7555-65.

47. Xiao ZY, Shangguan DH, Cao ZH, *et al.* Cell-specific internalization study of an aptamer from whole cell selection. Chemistry-a European Journal. [Article]. 2008;14(6):1769-75.

48. Kaur G, Roy I. Therapeutic applications of aptamers. Expert Opin Investig Drugs2008 Jan;17(1):43-60.

49. Yan A, Ellington A. A Hitchhiker's guide to HIV. Mol Ther2008 Aug;16(8):1356-8.