THE PROTEIN SOCIETY

A fusion-protein approach enabling mammalian cell production of tumor targeting protein domains for therapeutic development

Jia Hu,[1,2] Xiang Chen,[2]* Xuhua Zhang,[2,3] Xiaopeng Yuan,[2,4] Mingjuan Yang,[2] Hui Dai,[5] Wei Yang,[2] Qinghua Zhou,[1] Weihong Wen,[6] Qirui Wang,[2,7] Weijun Qin,[8] and Aizhi Zhao[2]*

[1]Lung Cancer Research Center, West China Hospital, Sichuan University, Chengdu, China
[2]Abramson Cancer Center, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania
[3]School of Life Sciences, Zhengzhou University, Zhengzhou, China
[4]Zhujiang Hospital, SouthernMedical University, Guangzhou, China
[5]Xinjiang Karamay Central Hospital, Karamay City, Xinjiang, China
[6]State Key Laboratory of Cancer Biology, Department of Immunology, Xijing Hospital, Fourth Military Medical University, ChinaXi'an
[7]College of Traditional Chinese Medicine, SouthernMedical University, Guang Dong, China
[8]Department of Urology, Xijing Hospital, Fourth Military Medical University, ChinaXi'an

Received 23 June 2017; Accepted 1 March 2018
DOI: 10.1002/pro.3399
Published online 3 March 2018 proteinscience.org

Abstract: A single chain Fv fragment (scFv) is a fusion of the variable regions of heavy (V_H) and light (V_L) chains of immunoglobulins. They are important elements of chimeric antigen receptors for cancer therapy. We sought to produce a panel of 16 extracellular protein domains of tumor markers for use in scFv yeast library screenings. A series of vectors comprising various combinations of expression elements was made, but expression was unpredictable and more than half of the protein domains could not be produced using any of the constructs. Here we describe a novel fusion expression system based on mouse TEM7 (tumor endothelial marker 7), which could facilitate protein expression. With this approach we could produce all but one of the tumor marker domains that could not otherwise be expressed. In addition, we demonstrated that the tumor associated antigen hFZD10 produced as a fusion protein with mTEM7 could be used to enrich scFv antibodies from a yeast display library. Collectively our study demonstrates the potential of specific fusion proteins based on mTEM7 in enabling mammalian cell production of tumor targeting protein domains for therapeutic development.

Keywords: recombinant protein expression; scFv; mammalian cells; cancer; yeast library screening; antibodies

Additional Supporting Information may be found in the online version of this article.

Grant sponsor: NIH (Transformative grant); Grant number: R01CA156695; Grant sponsor: National Natural Science Foundation of China; Grant number: 81502590; Grant sponsor: Honorable Tina Brozman Foundation, Claneil Foundation, and NCI FCCC-UPenn Cancer SPORE.

*__Correspondence to__: Xiang Chen, Abramson Cancer Center, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania. E-mail: chexmed@yahoo.com or Aizhi Zhao, Abramson Cancer Center, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania. E-mail: zhaoaizhi@gmail.com

Introduction

Recombinant protein expression systems are an important tool in both basic biological and applied medical research, as well as for commercial purposes. Although proteins are typically most easily expressed under physiological conditions within their native cell type, only a small fraction of proteins are naturally produced in quantities sufficient for study, or are secreted so that they can be readily purified. Commonly used expression systems include bacterial, yeast, insect, and mammalian cells. Important considerations in choosing the system are overall cost, cell density and protein yield, as well as ease and speed of protein production and purification.[1–4] Protein expression is usually more difficult to establish in mammalian cells because the culture of the cells is more expensive and complicated, and yields lower than that of microbial ones.[5] Target proteins of human origin, however, sometimes require post-translational modifications (e.g., *N*-glycosylation) for proper folding and biological activity which cannot be recapitulated in other cell types;[1–4,6,7] thus necessitating their production in mammalian cells. Yeast expression systems, for example, add sugar side chains with high mannose content,[8] and the glycosylation patterns for insect cell lines differ significantly from mammalian ones.[9] Indeed, due to the importance of post-translational modifications for protein function, cultivated mammalian cells are the dominant approach for producing recombinant proteins for clinical applications.[10,11] Therefore, the further development and optimization of mammalian cell expression systems is a critical field of study.

The secretory pathway of recombinant proteins in mammalian cells has been well characterized. Briefly, the plasmid DNA is transfected into the nucleus and the gene encoding the protein of interest is transcribed. Subsequently, messenger RNA translation begins in the cytoplasm and the emerging protein, directed by its signal peptide (SP), is translocated into the endoplasmic reticulum (ER) where it is folded and glycosylation is initiated.[12] The protein is then transported via a carrier vesicle to the golgi apparatus where it can be further modified[10] and sorted into secretory vesicles for transport to the cell membrane and release into the extracellular environment. Each of these steps represents a potential hurdle or blockade to protein expression. The secondary structure of the plasmid DNA, for example, can limit the level of transcription initiation, while misfolding within the ER can shut down production of the protein.[5] Moreover, growth conditions such as the medium used and temperature can influence protein yield.

Antibodies (Abs) are emerging as an important therapeutic approach in the treatment of cancer patients.[13] Abs can be designed to target unique or preferentially expressed epitopes on the surface of the tumor itself, or its microenvironment such as the vascular endothelium. The therapeutic effect of the Abs is achieved by manipulating tumor-related signaling or by promoting antitumor immune responses.[10,14] Payloads delivered by the Abs can include a range of cytotoxic agents[15] or immunocytokines.[16] We are interested in the development of therapeutic Abs through yeast antibody (scFv) library screenings with recombinantly produced extracellular domains of proteins, uniquely or preferentially, expressed on tumor cells and their vascular endothelium, which we collectively refer to as tumor marker extracellular domains (TMEDs). Many TMEDs have extensive post-translational modifications[17] and indeed all currently targeted in the clinic, such as HER-2, prostate-specific antigen, and carcinoembryonic antigen, are glycoproteins.[18,19] Here we sought to develop an optimized plasmid or set of plasmids for the mammalian cell expression of a panel of 16 TMEDs. Vector modifications had minimal impact on protein secretion levels and more than half of the TMEDs could not be synthesized at all. Therefore, under the assumption that these truncated proteins may be unstable in the ER, or unable to transit along the secretory pathway on their own, we created fusions of them with mouse TEM7protein, which itself is secreted at high levels. Remarkably, this strategy worked for all but one of the TMEDs that could not otherwise be expressed. Moreover, we demonstrated that the 205 amino acid extracellular domain of tumor antigen FZD10, which is predominantly expressed on tumor vascular endothelium in ovarian cancers,[20] and comprises three glycosylations and five disulfide bonds,[21] could be used to enrich a yeast library for scFv specificities that recognized cells expressing the native full-length, seven transmembrane FZD10 protein. Thus, we propose the use of mTEM7 fusion proteins as a general strategy for the reliable expression of extracellular protein domains for the purpose of developing Ab therapeutics against cancer and other diseases.

Results

Evaluation of TMED secretion by vector variants of pN1, pCDNA3.3, and pTT28

We began from cloning 16 human and mouse TMEDs including hTEM1, mTEM1, hTEM5, mTEM5, hTEM7, mTEM7, hCDCP1, mCDCP1, hFZD10, mFZD10, hSP17, mSP17, hEDNRB, hFAP, mFAP, and hFR (described in Materials and Methods and Supporting Information Table S1B) into a series of mammalian cell expression vectors (illustrated in Fig. 1). Of the 16 TMEDs, only six of them, hTEM1, mTEM1, hTEM7, mTEM7, hCDCP1, and hFR, could be produced in HEK293 cells using our panel of constructs. Shown in Figure 2(A–D) are

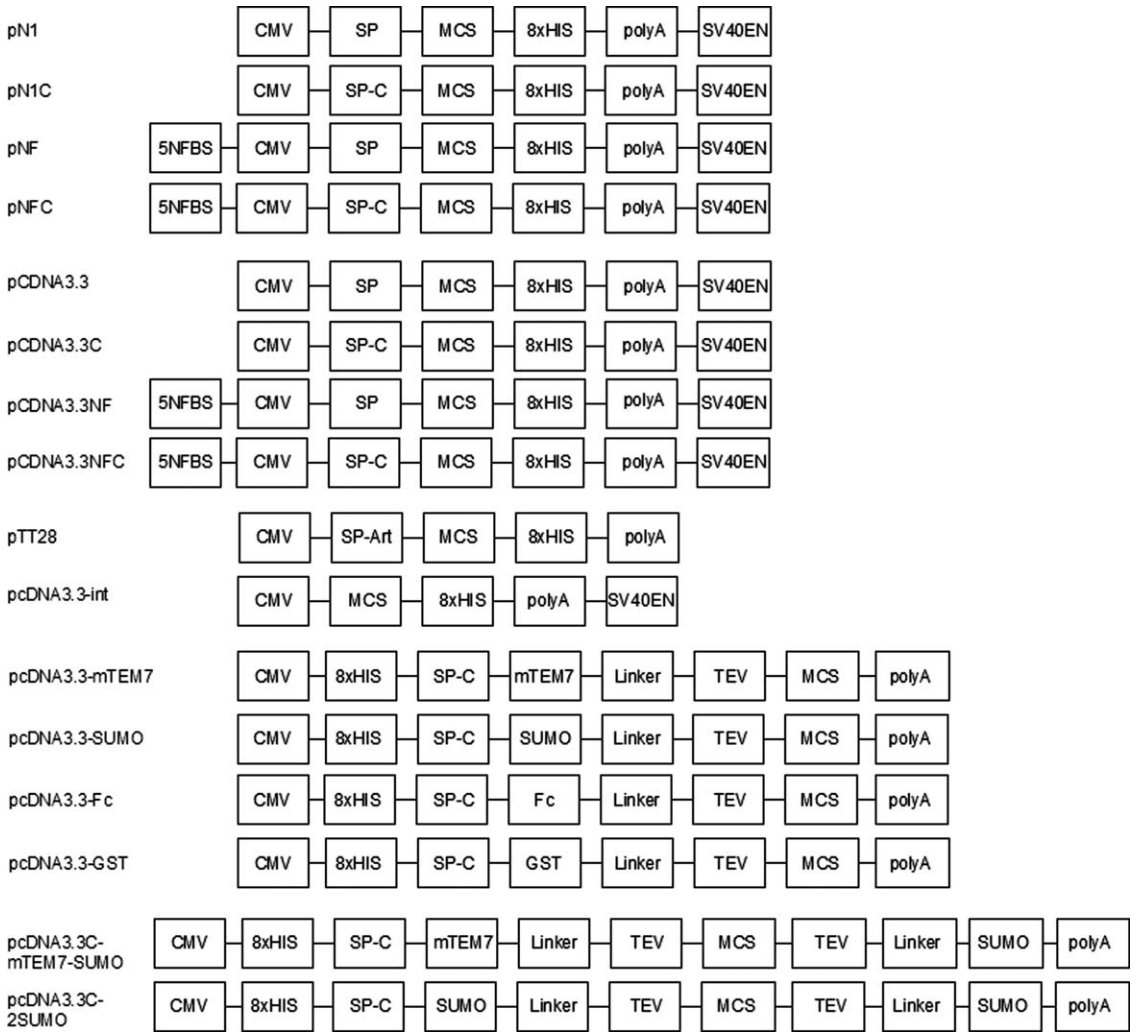

Figure 1. Illustration of mammalian cell expression vectors used to express TMEDs. Base vectors to express the TMEDs comprise pEGFP-N1 (abbreviated pN1), pcDNA3.3-TOPO, and pTT28. Elements incorporated into the various vector constructs include: five tandem repeats of the NF-κβ binding site (5NFBS), cytomegalovirus promoter (CMV), SP, codon-optimized SP (SP-C), artificial SP (SP-Art), multiple cloning site (MCS), eight-residue histidine tag (8 × HIS), polyadenine tail (polyA), TEV protease cleavage site (TEV), simian virus 40 enhancer (SV40EN), $(G_4S)_3$ linker (Linker), mutated human SUMO3 (SUMO), human immunoglobulin Fc fragment for IgG1 (Fc), and extracellular domain of mTEM7 (mTEM7).

representative Western blot analyses for hTEM1, mTEM1, hCDCP1, and hFR. For hTEM1 [Fig. 2(A)], the highest, and more or less equivalent, levels of expression were achieved for plasmids pcDNA3.3, pcDNA3.3C, and pcDNA3.3NF. Thus, neither codon optimization of the SP nor the addition of five tandem NF-*kb* binding sites (and the addition of TNF-α) improved yields. Moreover, hTEM1 expression by vector pcDNA3NFC, comprising both a codon optimized SP and the five tandem NF-*kb* binding sites, completely failed both in the presence and absence of TNF-α stimulation. However, a slight increase in hTEM1 secretion was observed upon codon optimization in the case of vector construct pN1C relative to pN1. In contrast, similar protein secretion levels of each mTEM1 and hFR were achieved for all tested vector constructs [Fig. 2(B,D)]. Lastly, of the vectors tested for hCDCP1, secretion was highest for vector pTT28 [Fig. 2(C)], which has neither the SV40 enhancer, nor the five tandem NF-*kb* binding sites, and comprises an artificial SP. Therefore, in general, protein expression levels were highly unpredictable, there was no obvious advantage conferred by any of the enhancing elements tested, and more than half of the TMEDs could not be produced.

Evaluation of intracellular protein expression by pcDNA3.3

Since we used the non-native SP VH3–20 for each of the TMEDs, we decided to remove it for the 10 proteins that failed to be secreted by any of the vectors tested, including hTEM5, mTEM5, mCDCP1, hFZD10, mFZD10,hSP17, mSP17, hEDNRB, hFAP, and, mFAP, and evaluated their production intracellularly. Amongst the 10 protein domains, only hSP17 and mSP17could be expressed intracellularly [Fig.

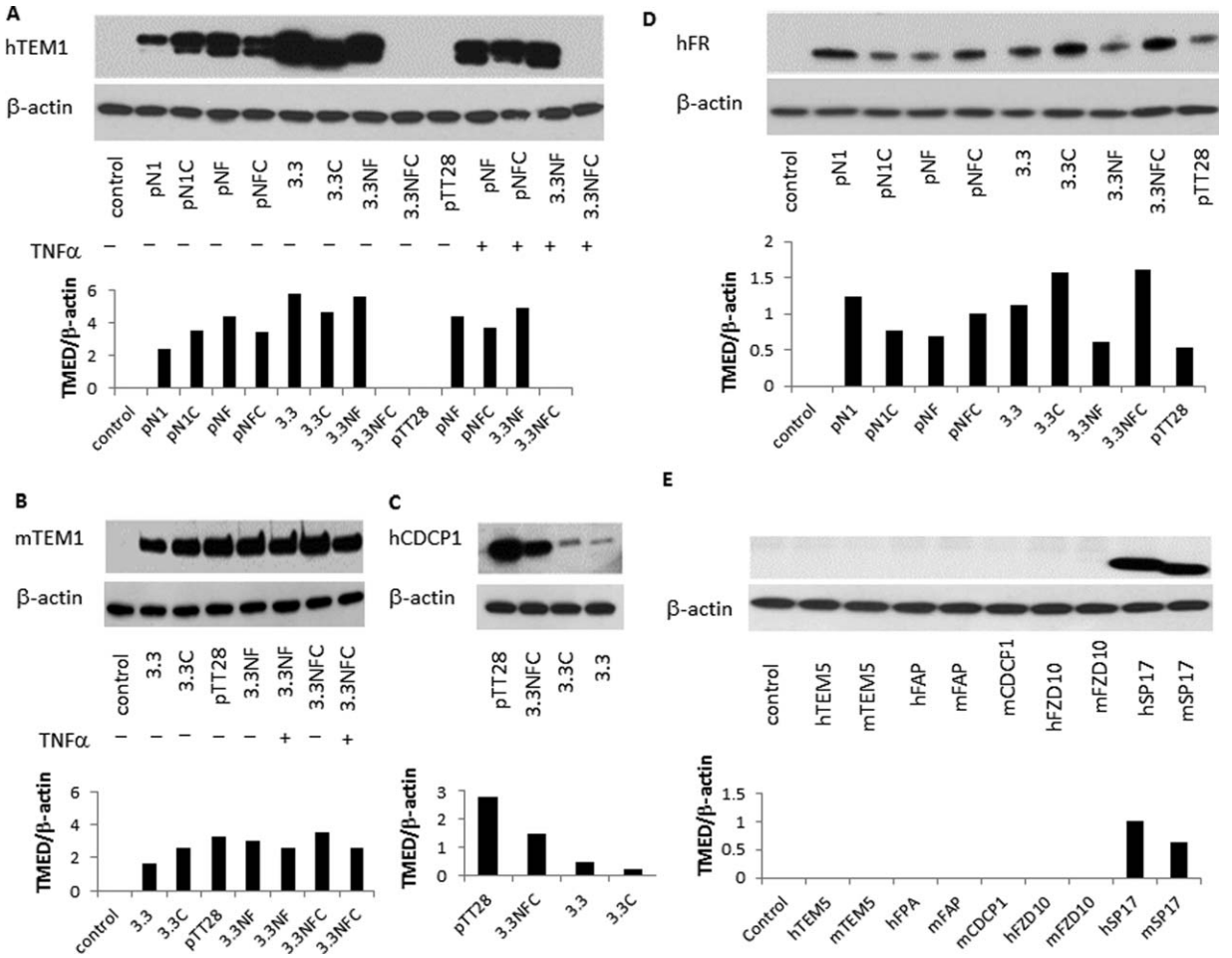

Figure 2. Expression of TMEDs by vector pTT28 and variants of pN1 and pcDNA3.3C. Shown by Western blot analysis with anti-HIS-HRP Ab are hTEM1 (A), mTEM1 (B), hCDCP1 (C), and folate receptor (D) found in serum free supernatant at 72 hr post-transfection of HEK293 cells. Protein expression was also evaluated for hTEM1 and mTEM1 expression following TNFα stimulation (A & B).Western blot analysis with anti-HIS-HRP Ab to evaluate intracellular expression of TMEDs, hTEM5, mTEM5, hFAP, mFAP, mCDCP1, hFZD10, mFZD10, hSP17, and mSP17 at 72 hr post-transfection (E). As a control for all Western blots, equal quantities of protein from the lysed cells were analyzed for β-actin expression. Band signal was quantified using ImagingJ software by multiplying the band area and mean intensity, and the intensity ratio of TMED/β-actin is plotted.

2(E)], albeit at low levels. Notably, SP17 is predicted to be a perimembrane protein, lacking both a SP and a transmembrane domain, and its mechanism for membrane anchoring remains unclear.[22,23] Therefore, unlike the other TMEDs, SP17 may not natively be folded within the ER. Intracellular expression, however, is not ideal because purification steps are more extensive; the cells must first be lysed, and as evidenced by the Western blot analysis in Figure 2(E) with anti-HIS Ab, there can be cross-reactivity of purification resin with other intracellular proteins. In addition, intracellular yields are typically lower than that of an optimized secretory system.

Fusion proteins for the expression of TMEDs which failed to be expressed by common vectors

As described above, more than half of the TMEDs failed to be produced via the secretory, or the intracellular, pathway for protein expression in HEK293 cells. Fusion proteins have been shown to be efficient formations to facilitate protein expression by fusing target protein to GST, Fc or SUMO,[24–26] here we supposed mTEM7 as a novel carrier protein for fusion protein expression by comparing with Fc and SUMO, which were reported previously. GST was excluded from our study because we found that it was poorly secreted on its own [Fig. 3(A)]. As described in the Materials and Methods, all of the fusion protein variants were cloned into pcDNA3.3C. We firstly demonstrated that the carrier proteins were secreted at high levels on their own [Fig. 3(A)]. Remarkably, all but two of the TMEDs (mTEM5 and mFAP) that could be secreted on their own were successfully produced as fusion proteins with mTEM7, SUMO and Fc [Fig. 3(B–D), respectively]. Amongst the carriers, mTEM7 was overall the best assessed, with relatively higher expression and low background [Fig. 3(B–D)]. The fusion construct pcDNA3.3C-mTEM7 was thus further used to produce the fusion

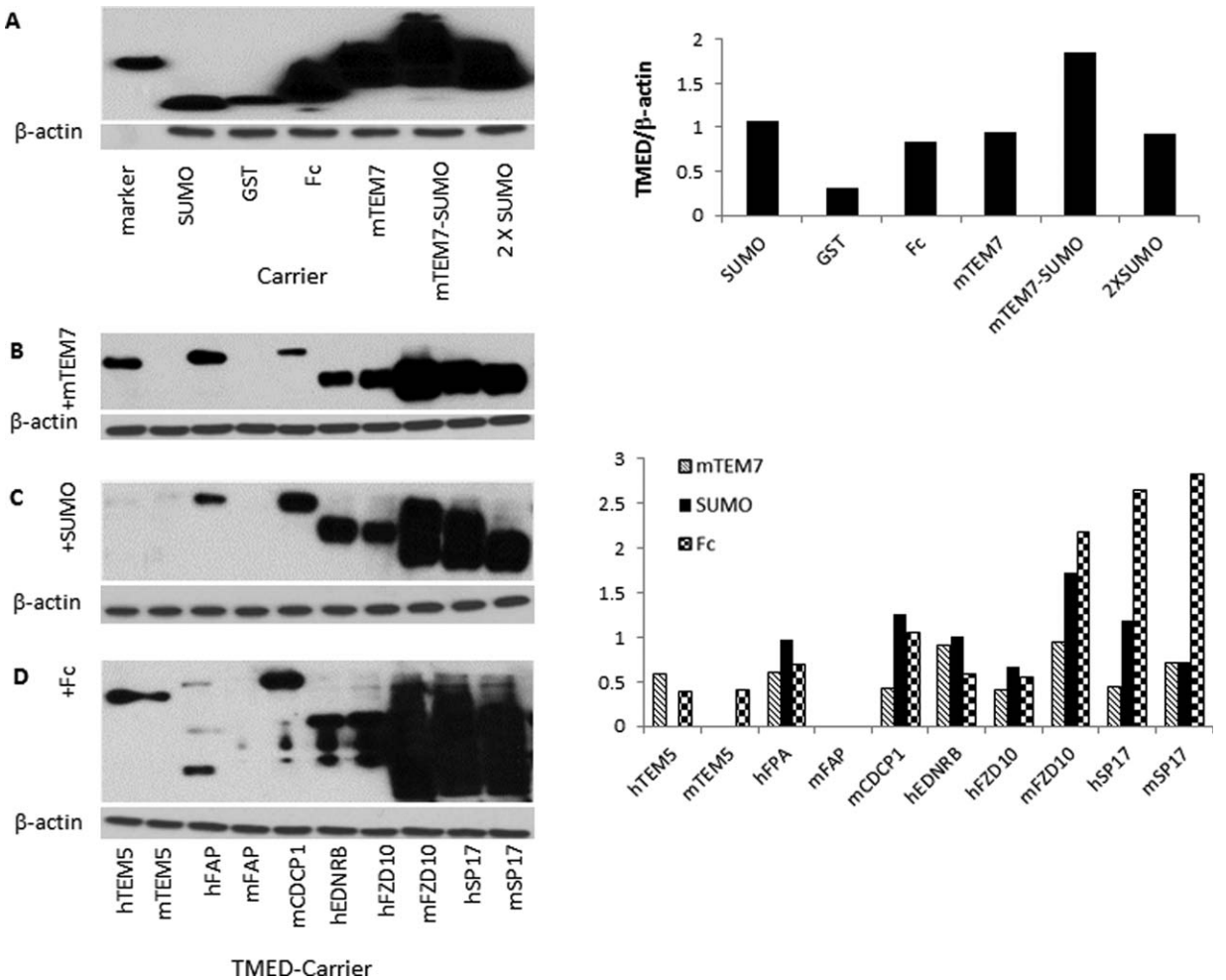

Figure 3. Production of difficult-expression TMEDs as fusion proteins in pcDNA3.3C with carriers SUMO, Fc mTEM7, SUMO-mTEM7, and 2 × SUMO. Shown by Western blot analysis with anti-HIS-HRP Ab is expression of the carrier proteins on their own (A). Also shown by Western blot analysis is expression of the TMEDs, hTEM5, mTEM5, hFAP, mFAP, mCDCP1, hEDNRB, hFZD10, mFZD10, hSP17, and mSP17with carrier proteins mTEM7 (B), SUMO (C), and Fc (D). Band signal was quantified using ImagingJ software by multiplying the band area and mean intensity, and the intensity ratio of carrier protein/β-actin (3A) or TMED-Carrier/β-actin (3B-D) is plotted.

proteins on a large-scale which were readily purified using Ni^{2+}-resin, with yields ranging between approximately 1 and 12 μg/mL (Supporting Information Table S2). TMEDs could be released from the mTEM7 by TEV protease digestion (data not shown). In a final attempt to produce mTEM5 and mFAP, we constructed two additional vectors in which the protein of interest was flanked by SUMO and mTEM7, or by 2 SUMOs (Fig. 1). Although the construct cDNA3.3C-SUMO-mTEM7 could not express mFAP, it did express mTEM5 (data not shown). The pcDNA3.3C-2SUMO vector, whereas, could express neither mTEM5 nor mFAP (data not shown).

scFv yeast library screening with recombinant fusion protein and evaluation of scFv cross-reactivity with native cell-surface expressed tumor markers

To evaluate the feasibility of using our fusion proteins for antibody isolation, we screened a yeast display scFv library with mTEM7-TMED fusion proteins, i.e., hFZD10, hTEM5, hFAP, and hSP17. Before doing this, however, we wished to evaluate steric hindrance of the carrier protein against potential epitopes at the interfacing amino-terminus region of the TMED. We thus incorporated three different tags, Flag, V5 and HA, at the carboxy-terminus of mTEM7 (the carrier) and by sandwich ELISA demonstrated their accessibility to Abs (Supporting Information Fig. S1). Prior to each round of yeast library screening, we negatively depleted the pools against the carrier protein, mTEM7, on its own. Three rounds of screening greatly enriched for yeast that could bind TMED but not the carrier (Fig. 4). In addition, the hFZD10-mTEM7 enriched yeasts were able to bind full-length hFZD10 expressed on the surface of the cell line MS1-hFZD10, while there was no binding to the MS1 cells on the their own (Fig. 5). Thus, mTEM7 not only enables the expression of difficult to produce

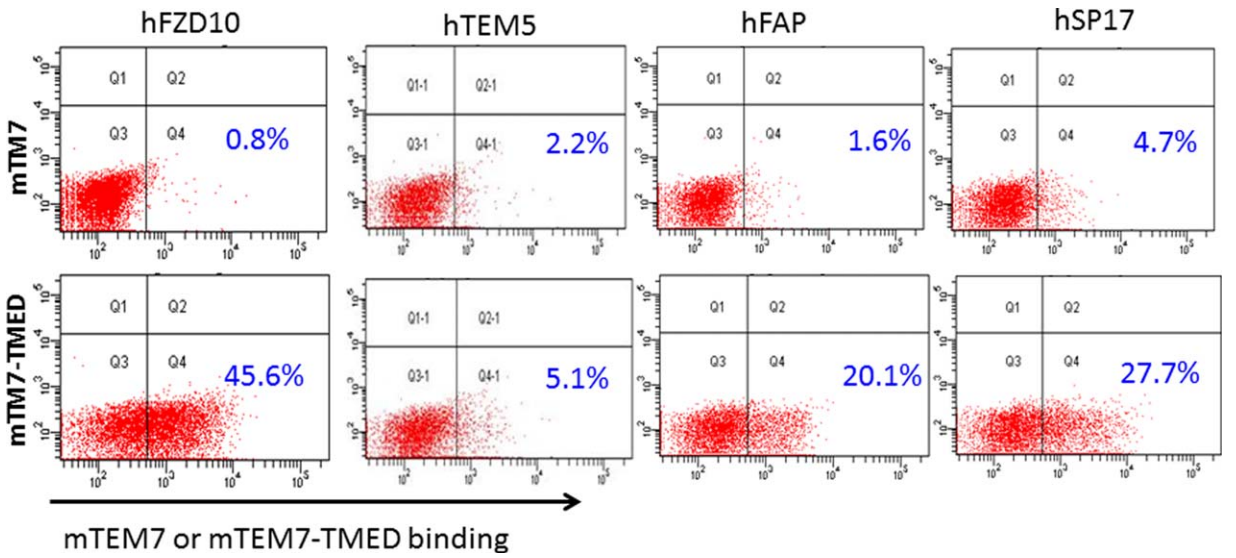

Figure 4. Enrichment of TMED-specific scFvs by yeast display scFv library screening with the fusion proteins. After three rounds of sorting with mTEM7-TMEDs and depletion against the carrier protein mTEM7, the resultant yeast populations were evaluated by flow cytometry for binding to biotinylated carrier protein mTEM7 or fusion protein TMED-mTEM7 (streptavidin-PE).

TMEDs, but they also enable the enrichment of scFv from yeast libraries that recognize the cell-surface full-length native protein, indicating correct folding of the recombinantly produced TMED.

ScFvs isolated from the MS1-hFZD10 binding yeasts also bind to the natively conformational hFZD10To further validate the fusion strategy for specific antibody generation for difficult-expression tumor markers, using hFZD10 as an example, we isolated 4 scFvs from the MS1-hFZD10 specific binding yeasts and checked their binding to mammalian cell surface expressed hFZD10 (MS1-hFZD10) that has a native conformation. The 4 scFvs were cloned into pcDNA3.3-Fc and expressed using HEK293 cells. All the four purified scFv-Fc recombinant proteins showed specific and significant binding to

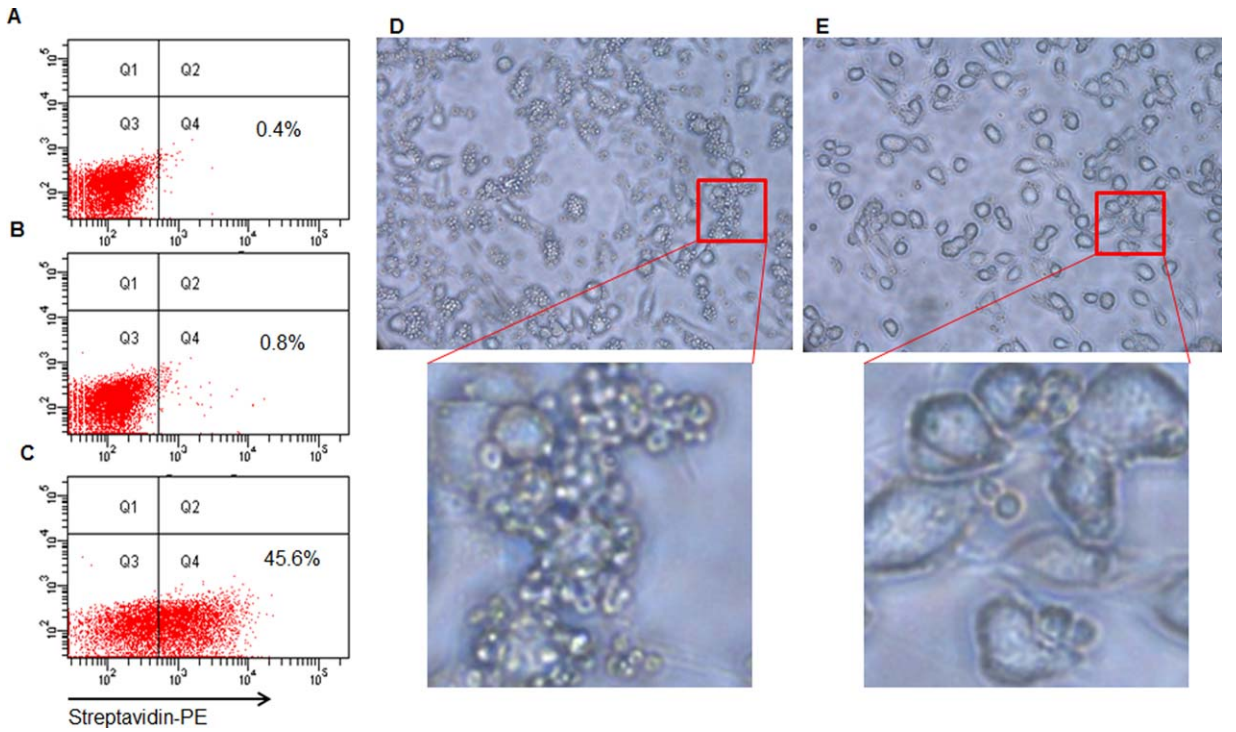

Figure 5. hFZD10-mTEM7 enriched yeast population also binds to the full-length hFZD10 expressed on cell surface. Comparing with negative control (A), flow cytometry showed hFZD10-mTEM7 enriched yeast population bind to MS1 cells (B), and MS1-hFZD10 cells (C). Representative microscope pictures showed these bindings on MS1-hFZD10 cells (D) and MS1 cells (E).

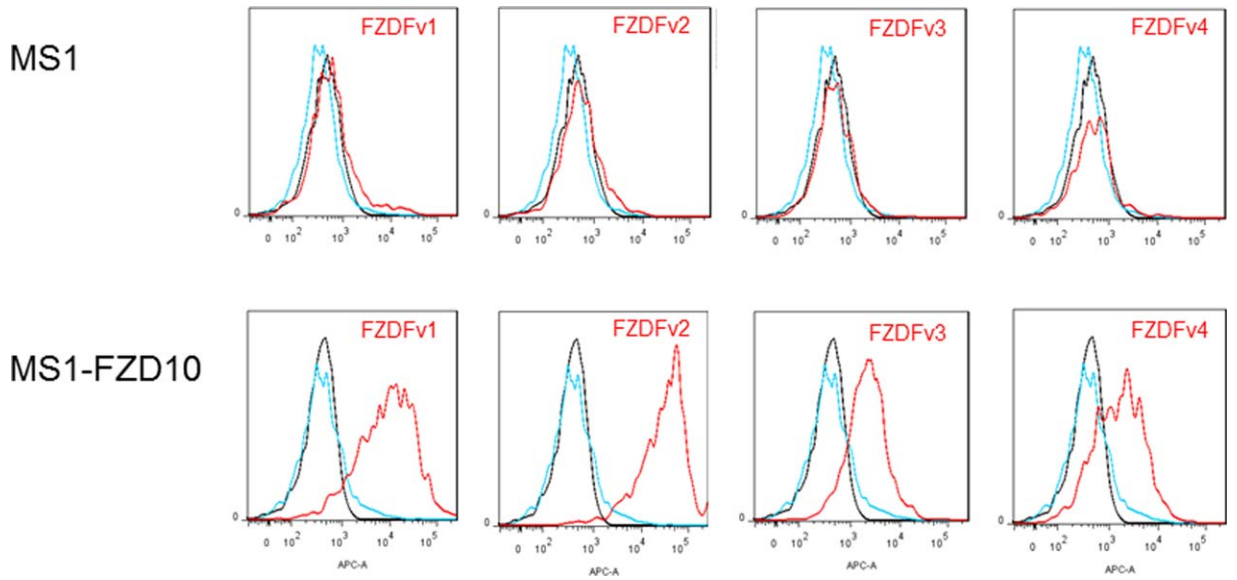

Figure 6. ScFvs isolated using TEM7-hFZD10 recombinant protein also bind to native conformation hFZD10 expressed on mammalian cell surface. The four anti-hFZD10 scFvs (FZDFv1, FZDFv2, FZDFv3, and FZDFv4) were prepared as fusion proteins to Fc and their binding to MS1 vs. MS1-hFZD10 was analyzed by flow cytometry. Black: untreated cells; cyanine blue: cells stained by the secondary Ab (APC-conjugated affinity pure F(ab')2 fragment goat anti-human IgG); red: cells stained sequentially with anti-hFZD10 Fc-scFvs and APC conjugated secondary Ab.

hFZD10 positive cell line MS1-hFZD10 but not hFZD10 negative cell line MS1 (Fig. 6). This indicated that TEM7-hFZD10 recombinant protein enriched yeast population display hFZD10-targeting scFvs that also recognize full length, native conformation hFZD10 expressed on mammalian cell surface and thereby mediate the specific binding of the yeast cells to MS1-hFZD10.

Discussion

The production of recombinant proteins in mammalian cells is influenced by complex molecular events occurring at the level of transcription, posttranscriptional processing, translation, posttranslational processing and secretion. Although the establishment and optimization of protein expression systems can consequently be challenging,[7,27] such efforts are of great value to research and development because many proteins, such as those upregulated or uniquely expressed by cancerous or other diseased cells, have important therapeutic potential.[28] Here we sought to develop an optimal plasmid or set of plasmids for the production, by transient gene expression in HEK293 cells, of a panel of 16 extracellular domains of tumor-associated surface proteins. Although protein yields are typically low for transient gene expression, as compared to the more time-consuming development of stable cell lines, this was not a concern because our ultimate purpose was to use the TMEDs in scFv yeast library screenings for the development of tumor-targeting Abs (i.e., relatively small quantities of the proteins would be required for the screenings).

We sought to optimize vector constructs for producing the different TMEDs in a mammalian cell secretory expression system for later use in the development of therapeutic Abs by scFv yeast library screening. We started with the base vectors pEGFP-N1, pcDNA3.3-TOPO and pTT28 (described in Materials and Methods). Features incorporated into the vectors included, (i) the SV40 enhancer, (ii) five tandem repeats of the NF-κβ binding site, and (iii) a codon-optimized VH3–20 SP. The SV40 enhancer was chosen because it can stimulate transcription from all mammalian promoters, regardless of orientation and up to thousands of base pairs away from the start site.[24,29,30] The NF-κβ binding site is a short nucleotide sequence which functions to enhance nuclear localization and expression of plasmid DNA. Upon cellular activation by TNF-α, NF-κβ that is normally sequestered in the cytosol will be released from its inhibitor proteins and functions to carry the plasmid into the nucleus[31–33] where it can also serve to enhance expression.[34] HEK293 cells have basal NF-κβ activity[35] and thus are a suitable cell-type for employing NF-κβ binding sites. The commonly used SP VH3–20 was used to direct the proteins along the secretory pathway. To avoid possible interference to protein expression by DNA primary and/or secondary structure the SP was also balanced for GC content by codon optimization. Amongst the 16 TMEDs that we sought to produce, 10 of them could not be expressed by any of the nine constructs. Moreover, of the six TMEDs that could be produced, the inclusion of expression enhancement elements had little if any positive

impact on protein yields. In fact, in the case of hTEM1, although vector pcDNA3.3NF yielded relative high levels of recombinant protein, codon optimization of the SP (pcDNA3.3NFC) led to a loss in protein production. In contrast, codon optimization has been reported by others to dramatically enhance protein expression in mammalian cells.[36–38] Although, we did not measure mRNA levels we speculate that this loss in expression may have resulted from a change in the secondary structure of the plasmid that was inhibitory to the initiation of transcription. Subsequent removal of the SP in vector pcDNA3.3 led to low intracellular expression of hSP17 and mSP17, but none of the other proteins could be produced along this pathway, presumably because they all require the distinct folding environment of the ER lumen.[12]

Under the assumption that the 10 TMEDs that could not be produced by any of our expression vectors was a result of improper folding, or faulty transit along the secretory pathway, we next attempted to express them as fusion proteins with the frequently used carrier protein Fc,[39,51–55] as well as the chaperone protein SUMO.[25,26] In addition, we assessed the ability of the well-secreted TMED, mTEM7, to serve as a carrier protein itself. Remarkably, the fusion protein approach enabled the production of all but two of the TMEDs, mTEM5 and mFAP, that could not otherwise be expressed. In a final attempt to express these domains we developed 'double' fusion protein constructs encoding mTEM7 and SUMO, or 2xSUMO, which enabled the production of mTEM5 but not mFAP. Yields for fusion and non-fusion TMEDs ranged from approximately 1 to 15 μg/mL. It should be noted, however, that no attempts were made to optimize culture conditions. Growing cells under hypothermic conditions, for example, which will modify cellular metabolism and increase transgene mRNA levels,[40] or culturing the cells in the presence of valproic acid, a histone deacetylase inhibitor,[41] could be explored to increase overall protein yield.

In summary, we have demonstrated that although different vectors may result in differences in the yield of a target protein, the modification of a given vector, such as the incorporation of tandem repeats of NF-κβ binding site, often has little to no effect on production levels. Moreover, it is impossible to predict *a priori* which if any proteins can be expressed using a given vector. However, for extracellular protein domains that cannot be produced on their own, either *via* secretory or intracellular protein pathways, we showed carrier proteins to be highly efficient tools. Overall we observed similar trends in protein production of TEMDs fused to different carrier proteins (i.e., TEMDs well expressed with one carrier protein were typically equally well expressed by others), and we found the novel carrier mTEM7 to enable the production of a superior quality of protein, as compared to Fc and SUMO, as observed by consistently cleaner bands on Western Blots. Critically, we successfully used the fusion protein mTEM7-hZFD10 to screen a yeast display scFv library and enrich for pools specific for the native, cell-surface expressed, transmembrane protein hZFD10. Thus, our study demonstrates the feasibility and potential of mTEM7 as a novel carrier protein for fusion proteins expression for the production of extracellular protein domains that cannot be secreted on their own, and these fusion proteins can be used in the enrichment of cross-reactive scFv having therapeutic potential.

Materials and Methods

Vector development

Constructs built in this study were developed from the commercially available mammalian expression vectors pEGFP-N1 (Clontech Laboratories Inc.) and pcDNA3.3-TOPO (Invitrogen) which both possess the SV40 enhancer. Five tandem repeats of the NF-κβ binding site were introduced upstream of the CMV promoter, and the most commonly used SP for directing proteins along the secretory pathway, VH3–20, was codon optimized for balanced GC content. The vector pTT28, comprising an artificial SP but neither the SV40 enhancer nor the NF-κβ binding site, was a gift from Dr. Xiaobo Duan at the BC Cancer Research Center in Vancouver, Canada.

Construction of vectors pN1 variants for secretory protein expression

An oligonucleotide (sequence 1) comprising the Kozak sequence,VH3–20 SP, a multiple cloning site (MCS), an 8xHis tag, and a stop codon, was created by overlap PCR using primers 1 and 2, digested with SpeI/NotI, and inserted into NheI/NotI digested pEGFP-N1 to replace the original MCS and the EGFP gene. The resulting vector was named pN1. Vector pNF was constructed by inserting five repeats of the NF-κβ binding site, prepared by annealing primers 3 and 4, upstream of the CMV promoter in pN1 vector digested with PciI. The GC content of the SP in pN1 and pNF was optimized (sequence-2) by amplifying the CMV-SP sequence using primers 5 and 6, and cloning it back into both pN1 and pNF by AseI/NheI digestion to create vectors pN1C and pNFC, respectively [Fig. 1, Supporting Information Table S1(A)].

Construction of vectors pcDNA3.3 variants for secretory protein expression

Primers 7 and 8 were used to amplify sequence-1 from pN1, which was then inserted into pcDNA3.3 TOPO by TA cloning to create pcDNA3.3. The five-repeat NF-κβbinding site was created by annealing

primers 9 and 10 and was inserted into vector pcDNA 3.3 by MluI single digestion to create vector pcDNA3.3NF. The SP sequence in pcDNA3.3 and pcDNA3.3NF was also codon-optimized by amplifying the CMV-SP sequence from vector pcDNA3.3 using primers 6 and 11 and cloned back into vector pcDNA3.3 and pcDNA3.3NF by NdeI/NheI digestion. The resulting vectors were named pcDNA3.3C and pcDNA3.3NFC [Supporting Information Table S1(A)].

Cloning of TMEDs into the pN1 and pcDNA3.3 expression vector variants

A panel of 16 human (h) and mouse (m) TMEDs were cloned into expression vectors. The TMEDs produced included: (i) the tumor endothelial markers hTEM1, mTEM1, hTEM5, mTEM5, hTEM7, mTEM7 which are elevated during tumor angiogenesis[42] (ii) the CUB domain containing protein 1, hCDCP1, and mCDCP1, which is preferentially expressed on highly metastatic human epidermoid carcinoma cells[43] (iii) frizzled homolog 10, hFZD10 and mFZD10 which is expressed on tumor vascular endothelium in ovarian cancers[20] (iv) the zona pellucida-binding protein, hSP17 and mSP17, which is involved in the migration of transformed cells[44] (v) enodothelin receptor type B, hEDNRB, which is implicated in enhancing angiogenesis in melanoma[45] and (vi) the fibroblast activation protein alpha, hFAP and mFAP, which is thought to contribute to invasiveness in malignant cancers[46] and (vii) human folate receptor, hFR, which is overexpressed by a number of epithelial derived tumors including ovarian, breast, renal, lung, colorectal and brain.[47] The cDNA for these genes were purchased from Open biosystems (ThermoScientific), except for hTEM7 and hFAP which were purchased from Origene, and hFZD10 which was synthesized with codon optimization to lower GC content and remove rare codons by Genscript. The precise residues for each of these TMEDs, as well as the primers used for their amplification are listed in Supporting Information Table 1B. The sequences of all constructs were confirmed by the facility of the Children's hospital of Philadelphia.

Construction of vector pcDNA3.3-intr for intracellular protein expression

Vector pcDNA3.3C was modified for intracellular recombinant protein expression by removal of the SP by XbaI/NheI digestion and insertion of the primers 12 and 13 annealed fragment, comprising a Kozak sequence and an ATG start codon. The resulting vector was named pcDNA3.3-intr. The TMEDs were amplified and cloned into this vector [Supporting Information Table S1(C)].

Construction of vectors pcDNA3.3C-SUMO, pcDNA3.3C-fc, and pcDNA3.3C-GST for secretory fusion protein expression

For enhanced expression, hSUMO3 was mutated (R58T, R60E)[26] by site-directed mutagenesis and then cloned into pcDNA3.3 to construct pcDNA3.3C-SUMO. Briefly, the amino-terminus of hSUMO3 was amplified from a cDNA template (Origene) using primers 14 and 15 to introduce a ClaI site and a 8xHis tag, and the carboxy-terminus of hSUMO3 was amplified using primers 16 and 17 to introduce a (G4S)3 linker and a TEV cleavage site (Glu-Asn-Leu-Tyr-Phe-Gln-Gly). Then the mutated (m) SUMO gene (encoding amino acids 2–92) was amplified by overlap PCR with primers 14 and 18 to introduce a multiple cloning site (MCS) downstream of SUMO. Finally the full fragment comprising 8xHis-mSUMO-linker-TEV-MCS was cloned into pcDNA3.3 by NheI/XhoI digestion of the vector and SpeI/XhoI digestion of the PCR product. The MCS includes restriction sites NheI, EcoRI, KpnI, SacII, HindIII, XhoI and BamHI. The resulting vector was named pcDNA3.3C-SUMO. To construct pcDNA3.3C-Fc and pcDNA3.3C-GST, genes encoding Fc and GST were amplified with primers 19 and 20, and primers 21 and 22, respectively, from their template in-house plasmids. The Fc and GST PCR products were digested with ClaI/HpaI and ligated into similarly digested pcDNA3.3C-SUMO (i.e., replacing SUMO). To avoid dimerization, the two cysteine residues at the hinge region of Fc were excluded. The TMEDs hTEM5, mTEM5, hFAP, mFAP, mCDCP1, hFZD10, mFZD10, hEDNRB, hSP17 and mSP17 were then amplified and cloned into these vectors using the indicated forward and reverse primers shown in Supporting Information Tables S1(A) and S1(B).

Construction of pcDNA3.3C-mTEM7 secretory fusion vector

Due to high secretion levels of mTEM7 in expression vector pcDNA3.3C, we assessed its ability to serve as a carrier protein. To create this fusion vector, primers 23 and 24 were annealed and inserted into NheI digested pcDNA3.3 that already harbored mTEM7 in order to introduce a 8xHis tag upstream of the TMED. Primers 26 and 27 were then annealed and inserted downstream of mTEM7 by digestion of the vector with XhoI to create a new (G4S)3 linker-MCS sequence including restriction sites NheI, ClaI, SacII, EcoRI, and XhoI. Subsequently, primers 27 and 28 were annealed and ligated into Nhe I/XhoI digested vector to introduce a TEV cleavage site[48,49] between the (G4S)3 linker and the MCS and thereby enable the TMED of interest to be released from its carrier protein post-secretion. The TMEDs hTEM5, mTEM5, hFAP, mFAP, mCDCP1, hFZD10, mFZD10, hEDNRB,

hSP17 and mSP17 were then amplified and cloned into this fusion vector.

FreeStyle 293F cell transfection

Expression of the proteins was assessed with the FreeStyle MAX System from Invitrogen that comprises a cationic lipid-based reagent that complexes with DNA for rapid transfection of HEK293 cells, mostly via endocytosis of the liposomes.[50] HEK293 cells are a preferred choice for transient gene expression because they can be grown under suspension in serum free media, and, like CHO cells, they have been approved for therapeutic protein production.[1] FreeStyle 293F cells were maintained in FreeStyle 293 expression medium and transient transfection of the cells conducted following the manufacturer's protocol (Invitrogen). Briefly, 37.5 μg plasmid DNA and 37.5 μL FreeStyle MAX reagent were diluted in 0.6 mL OptiPro SFM medium, respectively. After 10 min incubation the DNA-MAX complex was slowly added dropwise into 30 mL FreeStyle 293 cells at a density of 1×10^6/mL. For small-scale protein production, plasmid quantities and reagent volumes were proportionately scaled down for the transfection of 1 mL cell cultures (i.e., 1×10^6 cells) in 12-well plates.

Protein purification and Western blot analysis

Supernatants of the cultivated cells were collected 72 h post-transfection and the cells resuspended in an equal volume of fresh medium for a further 3 days of expression. The harvested supernatants (i.e., 60 mL per sample) were concentrated around 20-fold by centrifugation using Amicon ultra concentration tubes, diluted with equilibration buffer (0.05*M* phosphate buffered saline, 0.3*M* NaCl, pH 8.0), and then incubated with Ni^{2+} resin (Sigma) as per the manufacturer's suggestion for 2h at 4°C. Subsequently, the resin was washed three times with washing buffer (equilibration buffer containing 10 m*M* imidazole) and the protein was eluted with elution buffer (equilibration buffer containing 250 m*M* imidazole). Protein expression was detected in the 72 h supernatants by Western blot. Briefly, 10 μL supernatant was separated by SDS-PAGE gel, transferred onto PVDF membrane (Millipore) and detected by anti-His-HRP antibody (Santa Cruz Biotechnology Inc.).

Yeast scFv library screening with fusion protein TMED-mTEM7

The fusion protein hFZD10-mTEM7, hTEM5-mTEM7, hFAP-mTEM7, and hSP17-mTEM7 were used to screen a scFv yeast display library as previously described[28] using magnetic sorting combined with extensive depletion with the carrier protein mTEM7. Following three rounds of screening, binding of the enriched yeast-displayed scFvs to TMED-mTEM7 and mTEM7 was analyzed by flow cytometry. Briefly, the yeast cells were incubated with biotinylated proteins and binding detected by streptavidin-PE. For hFZD10, we have also established full-length hFZD10 expression cell line by transducing mouse endothelial cell line MS1 (MILE SVEN 1, (ATCC® CRL-2279™) C57BL/6) with full-length hFZD10 expression lentivirus and checked the binding of hFZD10-mTEM7 enriched yeast population to MS1 versus MS1-hFZD10 cells. Briefly, MS1 and hFZD10 expressing MS1 cells were grown on gelatin coated dishes and then incubated with yeast pool for 1 h at 4°C with rocking. The cells were then washed with PBS and binding was analyzed by microscopy and flow cytometry.

Identification and characterization of hFZD10 binding scFvs

From the MS-hFZD10 cell binding yeast population, four scFvs were identified using the method we described previously. The gene of the four scFvs were cloned into pcDNA3.3-Fc and expressed in HEK293F cells as fusion proteins with Fc at their N-terminus. The recombinant proteins were purified using nickel resin and their binding to the natively conformational hFZD10 expressed on MS1 cell surface were analyzed by flow cytometry. Briefly, MS1 and MS1-hFZD10 cells were washed twice with PBS, detached with 0.02% Versene buffer (1.37*M* NaCl, 26.8 m*M* KCl, 80.7 m*M* Na_2HPO_4, 14.7 m*M* KH_2PO_4, 5.4 m*M* disodium EDTA, 0.2% D-glucose), incubated with 100 n*M* anti-hFZD10 Fc-scFv at 4°C, washed three times with PBS and then incubated with 1:200 diluted APC-conjugated affinity pure F(ab')2 fragment goat anti-human IgG (Jackson Immunoresearch, PA) at 4°C. Finally cells were washed three times with PBS and Fc-scFv binding was analyzed by flow cytometry.

Author Contributions

J. H. and X. C. designed and performed the experiments and wrote a part of the manuscript; X. Z. performed part of experiments and wrote part of the manuscript; X. Y. prepared figures and wrote part of the manuscript; M. Y., H. D., and W. Y. performed part of experiments; Q. Z., Q. W., W. W., and W. Q. provided important reagents and/or plasmids. A. Z. directed the project and wrote the manuscript.

Conflict of Interest

The authors declare that they have no conflict of interest.

References

1. Baldi L, Hacker DL, Adam M, Wurm FM (2007) Recombinant protein production by large-scale transient gene expression in mammalian cells: state of the art and future perspectives. Biotechnol Lett 29:677–684.

2. Wurm F, Bernard A (1999) Large-scale transient expression in mammalian cells for recombinant protein production. Curr Opin Biotech 10:156–159.
3. Mariati Ho SC, Yap MG, Yang Y (2002) Post-transcriptional regulatory elements for enhancing transient gene expression levels in mammalian cells. Methods Mol Biol 801:125–135.
4. Lundstrom K (2010) Expression of mammalian membrane proteins in mammalian cells using Semliki Forest virus vectors. Methods Mol Biol 601:149–163.
5. Barnes LM, Bentley CM, Dickson AJ (2003) Stability of protein production from recombinant mammalian cells. Biotechnol Bioeng 81:631–639.
6. Ferrer-Miralles N, Domingo-Espín J, Corchero JL, Vázquez E, Villaverde A (2009) Microbial factories for recombinant pharmaceuticals. Microb Cell Fact 8:17.
7. Junge F, Schneider B, Reckel S, Schwarz D, Dötsch V, Bernhard F (2008) Large-scale production of functional membrane proteins. Cellular Mol Life Sci 65:1729–1755.
8. Gemmill TR, Trimble RB (1999) Overview of N- and O-linked oligosaccharide structures found in various yeast species. Biochim Biophys Acta 1426:227–237.
9. Jarvis DL, Kawar ZS, Hollister JR (1998) Engineering *N*-glycosylation pathways in the baculovirus-insect cell system. Curr Opin in Biotechnol 9:528–533.
10. Walsh G, Jefferis R (2006) Post-translational modifications in the context of therapeutic proteins. Nat Biotechnol 24:1241–1252.
11. Wurm FM (2004) Production of recombinant protein therapeutics in cultivated mammalian cells. Nat Biotechnol 22:1393–1398.
12. Braakman I, Bulleid NJ (2011) Protein folding and modification in the mammalian endoplasmic reticulum. Annu Rev Biochem 80:71–99.
13. Mellman I, Coukos G, Dranoff G (2011) Cancer immunotherapy comes of age. Nature 480:480–489.
14. Weiner LM, Surana R, Wang S (2010) Monoclonal antibodies: versatile platforms for cancer immunotherapy. Nat Rev Immunol 10:317–327.
15. Sievers EL, Senter PD (2013) Antibody-drug conjugates in cancer therapy. Annu Rev Med 64:15–29.
16. Spitaleri G, Berardi R, Pierantoni C, De Pas T, Noberasco C, Libbra C, González-Iglesias R, Giovannoni L, Tasciotti A, Neri D, Menssen HD, de Braud F (2013) Phase I/II study of the tumour-targeting human monoclonal antibody-cytokine fusion protein L19-TNF in patients with advanced solid tumours. J Cancer Res Clin Oncol 139:447–455.
17. Fuster MM, Esko JD (2005) The sweet and sour of cancer: glycans as novel therapeutic targets. Nat Rev Cancer 5:526–542.
18. Peracaula R, Barrabes S, Sarrats A, Rudd PM, de Llorens R (2008) Altered glycosylation in tumours focused to cancer diagnosis. Dis Markers 25:207–218.
19. Shariat SF, Karam JA, Margulis V, Karakiewicz PI (2008) New blood-based biomarkers for the diagnosis, staging and prognosis of prostate cancer. BJU Int 101:675–683.
20. Buckanovich RJ, Sasaroli D, O'Brien-Jenkins A, Botbyl J, Hammond R, Katsaros D, Sandaltzopoulos R, Liotta LA, Gimotty PA, Coukos G (2007) Tumor vascular proteins as biomarkers in ovarian cancer. J Clin Oncol 25: 852–861.
21. Koike J, Takagi A, Miwa T, Hirai M, Terada M, Katoh M (1999) Molecular cloning of Frizzled-10, a novel member of the Frizzled gene family. Biochem Biophys Res Commun 262:39–43.
22. Lea IA, Richardson RT, Widgren EE, O'Rand MG (1996) Cloning and sequencing of cDNAs encoding the human sperm protein, Sp17. Biochim Biophys Acta 1307:263–266.
23. Lim SH, Wang Z, Chiriva-Internati M, Xue Y (2001) Sperm protein 17 is a novel cancer-testis antigen in multiple myeloma. Blood 97:1508–1510.
24. Kumar R, Firak TA, Schroll CT, Subramanian KN (1986) Activation of gene expression is adversely affected at high multiplicities of linked simian virus 40 enhancer. Proc Natl Acad Sci USA 83:3199–3203.
25. Wang Z, Li H, Guan W, Ling H, Wang Z, Mu T, Shuler FD, Fang X (2010) Human SUMO fusion systems enhance protein expression and solubility. Protein Expr Purif 73:203–208.
26. Peroutka RJ, Elshourbagy N, Piech T, Butt TR (2008) Enhanced protein expression in mammalian cells using engineered SUMOfusions: Secreted phospholipaseA2. Protein Sci 17:1586–1595.
27. Bannwarth M, Schulz GE (2003) The expression of outer membrane proteins for crystallization. Biochim Biophys Acta 1610:37–45.
28. Zhao A, Nunez-Cruz S, Li C, Coukos G, Siegel DL, Scholler N (2011) Rapid isolation of high-affinity human antibodies against the tumor vascular marker Endosialin/TEM1, using a paired yeast-display/secretory scFv library platform. J Immunol Methods 363:221–232.
29. Byrne BJ, Davis MS, Yamaguchi J, Bergsma DJ, Subramanian KN (1983) Definition of the simian virus 40 early promoter region and demonstration of a host range bias in the enhancement effect of the simian virus 40 72-base-pair repeat. Proc Natl Acad Sci USA 80:721–725.
30. Moreau P, Hen R, Wasylyk B, Everett R, Gaub MP, Chambon P (1981) The SV40 72 base repair repeat has a striking effect on gene expression both in SV40 and other chimeric recombinants. Nucleic Acids Res 9: 6047–6068.
31. Fuseler JW, Merrill DM, Rogers JA, Grisham MB, Wolf RE (2006) Analysis and quantitation of NF-kappaB nuclear translocation in tumor necrosis factor alpha (TNF-alpha) activated vascular endothelial cells. Microsc Microanal 12:269–276.
32. Kikumori T, Kambe F, Nagaya T, Imai T, Funahashi H, Seo H (1998) Activation of transcriptionally active nuclear factor-kappaB by tumor necrosis factor-alpha and its inhibition by antioxidants in rat thyroid FRTL-5 cells. Endocrinology 139:1715–1722.
33. Fagerlund R, Kinnunen L, Kohler M, Julkunen I, Melen K (2005) NF-{kappa}B is transported into the nucleus by importin {alpha}3 and importin {alpha}4. J Biol Chem 280:15942–15951.
34. Mesika A, Grigoreva I, Zohar M, Reich Z (2001) A regulated, NFkappaB-assisted import of plasmid DNA into mammalian cell nuclei. Mol Ther 3:653–657.
35. Badr CE, Niers JM, Tjon-Kon-Fat LA, Noske DP, Wurdinger T, Tannous BA (2009) Real-time monitoring of nuclearfactor kappaB activity in cultured cells and in animal models. Mol Imaging 8:278–290.
36. Fath S, Bauer AP, Liss M, Spriestersbach A, Maertens B, Hahn P, Ludwig C, Schafer F, Graf M, Wagner R (2011) Multiparameter RNA and codonoptimization: a standardized tool to assess and enhance autologous mammalian gene expression. PLoS One 6:e17596.
37. Qiu JT, Chang TC, Lin CT, Chen YM, Li FQ, Soong YK, Lai CH (2007) Novel codon-optimized GM-CSF gene as an adjuvant to enhance the immunity of a DNA vaccine against HIV-1 Gag. Vaccine 25:253–263.
38. Kim CH, Oh Y, Lee TH (1997) Codon optimization for high-level expression of human erythropoietin (EPO) in mammalian cells. Gene 199:293–301.

39. Nilsson J, Ståhl S, Lundeberg J, Uhlén M, Nygren P (1997) Affinity fusion strategies for detection, purification, and immobilization of recombinant proteins. Protein Expr Purif 11:1–16.
40. Al-Fageeh MB, Marchant RJ, Carden MJ, Smales CM (2006) The cold-shock response in cultured mammalian cells: harnessing the response for the improvement of recombinant protein production. Biotechnol Bioeng 93: 829–835.
41. Backliwal G, Hildinger M, Kuettel I, Delegrange F, Hacker DL, Wurm FM (2008) Valproic acid: a viable alternative to sodium butyrate for enhancing protein expression in mammalian cell cultures. Biotechnol Bioeng 101:182–189.
42. Carson-Walter EB, Watkins DN, Nanda A, Vogelstein B, Kinzler KW, St Croix B (2001) Cell surface tumor endothelial markers are conserved in mice and humans. Cancer Res 61:6649–6655.
43. Hooper JD, Zijlstra A, Aimes RT, Liang H, Claassen GF, Tarin D, Testa JE, Quigley JP (2003) Subtractive immunization using highly metastatic human tumor cells identifies SIMA135/CDCP1, a 135 kDa cell surface phosphorylated glycoproteinantigen. Oncogene 22: 1783–1794.
44. Wen Y, Richardson RT, Widgren EE, O'Rand MG (2001) Characterization of Sp17: a ubiquitous three domain protein that binds heparin. Biochem J 357:25–31.
45. Lahav R, Suva ML, Rimoldi D, Patterson PH, Stamenkovic I (2004) Endothelin receptor B inhibition triggers apoptosis and enhances angiogenesis in melanomas. Cancer Res 64:8945–8953.
46. Ghersi G, Zhao Q, Salamone M, Yeh Y, Zucker S, Chen WT (2006) The protease complex consisting of dipeptidyl peptidase IV and seprase plays a role in the migration and invasion of human endothelial cells in collagenous matrices. Cancer Res 66:4652–4661.
47. Kelemen LE (2006) The role of folate receptor alpha in cancer development, progression and treatment: cause, consequence or innocent bystander? Int J Cancer 119: 243–250.
48. Waugh DS (2011) An overview of enzymatic reagents for the removal of affinity tags. Protein Exp Purif 80: 283–293.
49. Miladi B, Bouallagui H, Dridi C, El Marjou A, Boeuf G, Di Martino P, Dufour F, Elm'Selmi A (2011) A new tagged-TEV protease: Construction,optimisation of production, purification and test activity. Protein Expr Purif 75:75–82.
50. Douglas KL, Piccirillo CA, Tabrizian M (2008) Cell line-dependent internalization pathways and intracellular trafficking determine transfection efficiency of nanoparticle vectors. Eur J Pharm Biopharm 68:676–687.
51. Nakano H, Yamazaki T, Ikeda M, Masai H, Miyatake S, Saito T (1994) Purification of glutathione S-transferasefusion proteins as a non-degraded form by using a protease-negative *E. coli* strain, AD202. Nucleic Acids Res 22:543–544.
52. Smith DB, Johnson KS (1988) Single-step purification of polypeptides expressed in Escherichia coli as fusions with glutathione S-transferase. Gene 67:31–40.
53. Smith DB, Corcoran LM (2001) Expression and purification of glutathione-S-transferase fusion proteins. Curr Protoc Mol Biol Chapter 16:Unit16.7.
54. Peters RT, Low SC, Kamphaus GD, Dumont JA, Amari JV, Lu Q, Zarbis-Papastoitsis G, Reidy TJ, Merricks EP, Nichols TC, Bitonti AJ (2010) Prolonged activity of factor IX as a monomeric Fc fusion protein. Blood 115: 2057–2064.
55. Fisher CJ, Jr Agosti JM, Opal SM, Lowry SF, Balk RA, Sadoff JC, Abraham E, Schein RM, Benjamin E (1996) Treatment of septic shock with the tumor necrosis factor receptor: Fc fusion protein. N Engl J Med 334:1697–1702.