

# SCIENTIFIC REPORTS



OPEN

## Small Peptide Ligands for Targeting EGFR in Triple Negative Breast Cancer Cells

Hanieh Hossein-Nejad-Ariani, Emad Althagafi & Kamaljit Kaur

The efficacy of chemotherapy for cancer treatment can be increased by targeted drug delivery to the cancer cells. This is particularly important for triple negative breast cancer (TNBC) for which chemotherapy is a major form of treatment. Here we designed and screened a library of 30 peptides starting with a previously reported epidermal growth factor receptor (EGFR) targeting peptide GE11 (YHWYGYTPQNV). A direct peptide array-whole cell binding assay, where the peptides are conjugated to a cellulose membrane, was used to identify four peptides with enhanced binding to TNBC cells. Next, the four peptides were synthesized as FITC-labelled soluble peptides to study their direct uptake by TNBC cells using flow cytometry. The results showed that peptide analogue 22 had several fold higher uptake by the TNBC cells compared to the lead peptide GE11. The specific uptake of the peptide analogue 22 was confirmed by competition experiment using pure EGF protein. Further, peptide 22 showed dose dependent uptake by the TNBC MDA-MB-231 cells ( $10^5$ ) with uptake saturating at around  $2\ \mu\text{M}$  peptide concentration. Thus, peptide 22 is a promising EGFR specific TNBC cell binding peptide that can be conjugated directly to a chemotherapeutic drug or to nanoparticles for targeted drug delivery to enhance the efficacy of chemotherapy for TNBC treatment.

Triple negative breast cancer (TNBC) is an important subtype of breast cancer as it is more aggressive than other subtypes and has poor prognosis<sup>1–4</sup>. TNBC is characterized as a breast cancer phenotype that is estrogen receptor (ER) negative, progesterone receptor (PR) negative, and is without the overexpression of human epidermal growth factor receptor 2 (HER2), and hence is called the “triple-negative” phenotype<sup>2</sup>. TNBC is more common in younger women, premenopausal African American women, and women of lower socioeconomic status (SES) when compared to other subtypes. Diagnosis of TNBC usually happens at the more advanced stage of the disease and it usually leads to death within the first five years of diagnosis. Additionally, in TNBC, high grade rapidly growing tumors are poorly differentiated and metastasis tends to occur in soft tissues including lungs and brain rather than in bone<sup>4</sup>. Chemotherapy is a major way to treat TNBC. However it has been difficult to target chemotherapy to TNBC due to the lack of well-defined molecular targets<sup>3,5,6</sup>.

Peptide 1 (or GE11, Table 1) is a dodecapeptide that binds specifically to epidermal growth factor receptor (EGFR or ErbB1) overexpressed in a number of tumors of epithelial origin including breast cancer, and is being used as a cancer cell targeting peptide<sup>7–10</sup>. EGFR has been suggested to play an important role in TNBC<sup>11,12</sup>. Several TNBC cell lines, like MDA-MB-468 and MDA-MB-231, show high expression levels of EGFR<sup>13,14</sup>. Although therapeutic monoclonal antibodies like cetuximab and tyrosine kinase inhibitors like lapatinib have not shown much success in clinical trials for TNBC patients, several EGFR-targeted antibody-drug conjugates (ADCs) show great potential for TNBC treatment and are currently undergoing clinical trials for TNBC patients<sup>11</sup>. Similarly peptide-drug conjugates can be designed for targeting EGFR in TNBC. The cancer cell targeting peptides are not therapeutic like antibodies, but have several advantages as targeting ligands over antibodies such as they can be easily synthesized, are more economical, can be designed as stable analogues, and display low immunogenicity.

Several peptides that target breast cancer cells have been engineered for enhanced specificity and proteolytic stability for clinical applications<sup>15–17</sup>. Peptides like RGD<sup>18</sup>, NGR<sup>19–21</sup>, and p160 (VPWMEPAYQRFL)<sup>22–24</sup> target a number of different receptors such as integrin, aminopeptidase N (or CD13) and keratin 1, respectively, overexpressed by breast cancer cells. A cyclic analogue of peptide p160, N- to C terminal cyclized WXEAAYQkFL

Chapman University School of Pharmacy (CUSP), Harry and Diane Rinker Health Science Campus, Chapman University, Irvine, California, 92618-1908, USA. Correspondence and requests for materials should be addressed to K.K. (email: [kkaur@chapman.edu](mailto:kkaur@chapman.edu))

Peptide#	Sequence	Relative Cell Adhesion			
		MDA-MB-468	MDA-MB-231	MDA-MB-436	MCF-10A
1	YHWYGYTPQNVI	1.0	1.0	1.0	1.0
2	WQTNVIHPYVYG	1.1	1.2	1.1	1.0
3	YGPWYNHYITQV	1.0	1.1	1.1	1.0
4	VPWXPAYQRFL	1.1	1.3	1.1	1.2
5	WXEAAYQRFL	1.3	1.7	1.8	1.7
6	AHWYGYTPQNVI	1.2	1.3	1.0	1.1
7	YAWYGYTPQNVI	1.1	1.1	1.0	1.0
8	YHAYGYTPQNVI	0.8	1.0	0.9	0.9
9	YHWAGYTPQNVI	1.3	1.4	1.2	1.0
10	YHWYAYTPQNVI	1.3	1.5	1.2	1.1
11	YHWYGATPQNVI	1.2	1.3	1.0	1.0
12	YHWYGYAPQNVI	1.0	1.3	1.0	1.1
13	YHWYGYTAQNVI	1.3	1.5	1.1	1.1
14	YHWYGYTPANVI	1.4	1.7	1.2	1.2
15	YHWYGYTPQAVI	1.2	1.6	0.9	1.1
16	YHWYGYTPQNAI	1.1	1.4	1.0	1.2
17	YHWYGYTPQNV	1.4	1.6	1.2	1.2
18	YHWYGYTPQNV	1.4	1.9	1.3	1.2
19	YHWYGYTPQN	1.3	2.0	1.1	1.1
20	HWYGYTPQNVI	1.2	1.5	1.0	1.2
21	WYGYTPQNVI	1.4	1.9	1.0	1.2
22	YHWYGYTPENVI	1.6	2.6	1.3	1.2
23	YHWYGYTPQDVI	1.6	2.8	1.0	1.4
24	YKWYGYTPQNVI	1.1	1.4	1.0	1.0
25	YRWYGYTPQNVI	1.4	1.9	1.0	1.2
26	YHWYGYTPKNVI	1.3	1.6	1.0	0.8
27	YHWYGYTPQKVI	1.2	1.8	1.0	1.0
28	FHWYGYTPQNVI	1.2	1.5	0.8	1.2
29	YHWFYGYTPQNVI	1.2	1.7	1.1	1.1
30	YHWYGFYTPQNVI	1.3	1.5	0.9	1.0

**Table 1.** Designed peptide GE11 (or peptide 1) library and the relative cell adhesion ratio of the peptides. The letter X in the peptide sequence represents the amino acid norleucine (Nle).

(where k is D-lysine), was recently reported that preferentially accumulates in tumor after injection in mice carrying orthotopic breast MDA-MB-231 tumors<sup>25</sup>. Similarly, peptide 1 is a good candidate peptide for breast cancer targeting as it binds EGFR, and can be engineered for proteolytically stability and specific targeting of TNBC cells. Peptide 1 was identified by screening a phage display peptide library against purified human EGFR protein<sup>8</sup>. The peptide is internalized by cells overexpressing EGFR and when delivered i.v. in mice it accumulates in EGFR overexpressing tumor xenografts. Peptide 1 has been used as the targeting ligand for delivery of genes using poly-ethylenimine (PEI) polyplexes and antisense oligonucleotide nanoparticles<sup>8,26</sup>. A number of studies have explored peptide 1 as a targeting ligand in nanoparticles to treat cancers including breast and ovarian malignant cells<sup>27–31</sup>.

Here we have designed a synthetic peptide library of 29 analogues of peptide 1 (GE11)<sup>8</sup> with an aim to target EGFR specifically in TNBC cells (Table 1). The library was synthesized directly on cellulose membrane where the peptides remain covalently conjugated to the membrane. The library was screened against TNBC cells by incubating the membrane with different cells followed by labeling the bound cells with a fluorescent dye. Peptides showing highest relative fluorescence compared to the lead peptide (peptide 1) were selected from the whole cell-peptide binding assay, and were individually synthesized as well as their uptake in normal and TNBC cells was compared. Peptide 4 (or p160) reported previously to target breast cancer cells was used as a positive control peptide<sup>23–25</sup>. Several new analogues of peptide 1 with higher binding (using whole cell-peptide binding assay) and uptake (using flow cytometry) to TNBC were identified. The new analogues of peptide 1 that target EGFR and specifically bind TNBC cells can be used for targeted delivery of chemotherapeutics like doxorubicin for TNBC treatment.

## Results and Discussion

**Peptide Array Synthesis.** A peptide array consisting of 30 peptides in duplicates (60 spots) was synthesized on functionalized cellulose membrane (amino-PEG<sub>500</sub>) using automated spot synthesis ResPep SL apparatus. The free amino functional group present on the surface of the membrane was used to covalently couple the first amino acid, and thereafter the remaining amino acids were coupled sequentially to complete the sequences. Peptide array was synthesized using 96 spot mode (12 columns and 8 rows). Each peptide was synthesized in duplicates and a spot was skipped in order to separate from the next peptide sequence.

**Peptide Library.** Peptide 1 (GE11, Table 1) is a 12-mer peptide rich in hydrophobic aromatic amino acids (three tyrosines and one tryptophan). Using previous strategy<sup>21,32</sup>, we designed a library consisting of GE11-derived peptide sequences. The library of 30 peptides includes negative (2 and 3, scrambled) and positive (4 or peptide p160 and 5 or peptide 18 from a previous study)<sup>33</sup> control sequences. There were also sequences with alanine scan (peptides 6–17), C-terminus truncated sequences (18–19), N-terminus truncated sequences (20–21), and negatively (22–23) and positively (24–27) charged analogues. Finally the peptide library also had sequences with Y to F substitutions (28–30).

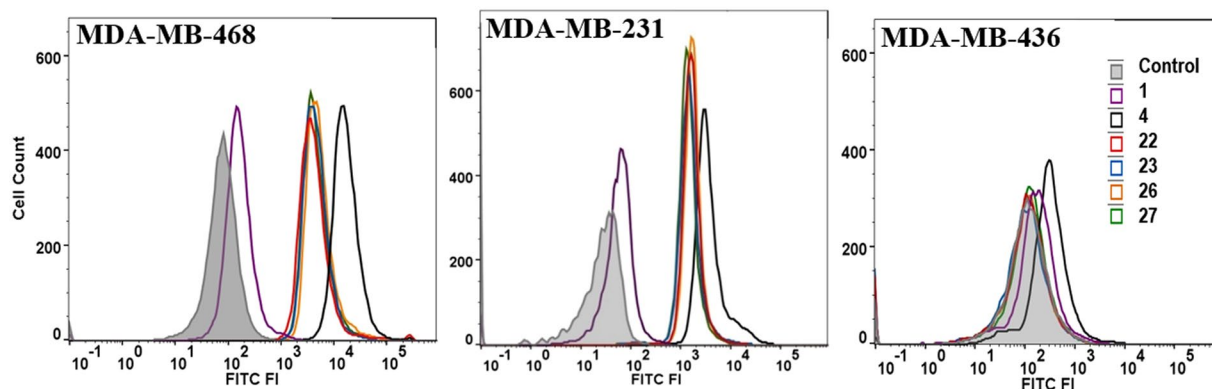
Peptide array on the cellulose membrane was used to evaluate the binding of each peptide to different cell lines. The amino functionalized cellulose membrane provided by Intavis is acid hardened providing high stability at pH range from 1–14, permitting its use in acidic or basic conditions while maintaining membrane integrity. The PEG linker on membrane surface functions as a hydrophilic spacer between the peptide and the solid surface (membrane) reducing background signal. Cell lines were carefully selected for screening peptides as our goal here is to screen peptides with high binding/uptake by the TNBC cells. Our lead peptide, peptide 1 (or GE11) targets and binds EGFR. EGFR is overexpressed by a number of TNBC cells such as MDA-MB-468 (obtained from primary tumor source) and MDA-MB-231 (from metastasis cancer tissue)<sup>13,14</sup>. While other TNBC cells like MDA-MB-436 (obtained from metastasis cancer tissue) are known to express low levels of EGFR. Non-tumorigenic epithelial MCF10A and MCF12A cells derived from the breast tissue also express low levels of EGFR<sup>14,34</sup>. We selected two TNBC cell lines with relatively high expression of EGFR expression (MDA-MB-468/231) and a TNBC cell line with low level of EGFR expression (MDA-MB-436) for the peptide array screening experiments. In addition, non-tumorigenic cell line, MCF-10A was used for comparison. These four cell lines were allowed to bind the membrane by seeding the cells on the membrane and the bound cells were labeled with a fluorescent dye CyQUANT.

**Peptide Library Screening.** To screen the library and identify high binding peptides to TNBC cells, the relative cell adhesion was monitored by comparing the fluorescence of cells (cancer or non-cancerous) bound to wild type peptide 1 (Table 1). Peptides 22 and 23 displayed highest affinity (up to 2.8-fold compared to peptide 1) to the three TNBC cell lines. These two peptides with Q9E and N10D substitutions, respectively, are negatively charged (net charge –1). Two other peptides that also showed high affinity were peptides 26 and 27. Here again the two same amino acids (Q9 and N10) were substituted, however, with positively charged lysine. Analysis of the alanine scan showed that substitution of these two residues (Q9 and N10) with alanine (especially peptide 14 with Q9A substitution) also gives higher binding to cancer cells. These results suggest that substitution of Q9 with a negatively charged glutamic acid (E) leads to highest cancer cell binding. In addition, alanine scan suggested that W3 is important for cancer cell binding as substitution of this residue with alanine led to decrease in relative cell adhesion ratio (0.8–1.0). The positive control peptide 5 showed higher binding to TNBC cells compared to peptide 1 (1.3–1.8-fold), however, positive control peptide 4 did not show much improvement in binding based on the peptide array-whole cell binding assay. Based on these results, four peptides 22, 23, 26, and 27 that displayed highest binding to the TNBC cells and did not show any substantial increase in affinity for the non-cancerous cells (MCF10A) were selected for cell uptake studies.

**In Vitro Cell Uptake of Soluble (Free) Peptides.** Flow cytometry was employed to determine the cellular uptake of selected six peptides (1, 4, 22, 23, 26, and 27). As opposed to the membrane bound (conjugated) peptide in library screening, the peptides used here were soluble (or free) FITC-labeled peptides. Peptides were labeled with FITC in the N-terminal and a  $\beta$ -alanine residue was used as a spacer between the FITC and the peptide sequence. The same three TNBC cells lines (MDA-MB-468, MDA-MB-231 and MDA-MB-436) used for peptide array screening were used here as well. Peptides 1 and 4 were used as controls. Peptide 1 is the lead EGFR binding peptide, whereas, peptide 4 binds keratin 1. Peptide 4 was included in this experiment to compare the uptake of this keratin 1 binding peptide with the EGFR binding peptides. Peptides 22, 23, 26, and 27 were identified as the best EGFR binding peptides from the library screening.

Figure 1 shows histograms and the mean fluorescence intensity (MFI) values for the three TNBC cell lines treated with six FITC-labeled peptides. Cell population treated with peptides 4, 22, 23, 26, and 27 showed significant FITC fluorescence in MDA-MB-468 and MDA-MB-231 when compared with cells treated with media only or treated with peptide 1. Peptide 4 showed much higher uptake (MFI 23741) compared to the EGFR binding peptides (MFI 8384 or less). Among the new EGFR binding peptides, peptide 22 showed the highest uptake by the TNBC cells. In addition, the uptake of peptide 22 to MDA-MB-468 and MDA-MB-231 cells improved substantially with 123-fold and 47-fold higher affinity, respectively, compared to peptide 1. The peptides showed highest uptake by the MDA-MB-468 cells followed by MDA-MB-231, and the lowest uptake was for MDA-MB-436 cells. Interestingly this is also the order of EGFR expression that is reported for these three cell lines with MDA-MB-468 showing highest level of EGFR expression<sup>13,14</sup>. The uptake of peptides 26 and 27 is only slightly less compared to peptide 22 even though these peptides have a substitution of a positively charged residue (K) for Q9 or N10 as opposed to negatively charged residue (E) for Q9 in peptide 22 (Fig. 1 table). It is anticipated that these peptide binds to the same site on EGFR as the native ligand EGF, and comparison of the amino acid sequence of these peptides and EGF shows that E and R may be preferred residues at positions 9 and 10 (vide infra).

In addition, it is notable that the uptake of the peptides by the cells is affected by the presence of free C-terminal carboxylic group, because the membrane bound peptides without free C-terminal carboxylic group showed less binding affinity to the cells compared to the soluble (or free) FITC labeled peptides which contain the free C-terminal carboxylic group. For instance peptide 22 showed only 2–3 fold improvement compared to peptide 1 (Table 1) using the peptide array-whole cell binding assay, whereas the uptake of peptide 22 by MDA-MB-468/231 cells is 50–123 fold compared to peptide 1 when evaluated using soluble (free) peptides *in vitro* uptake using flow cytometry (Fig. 1).



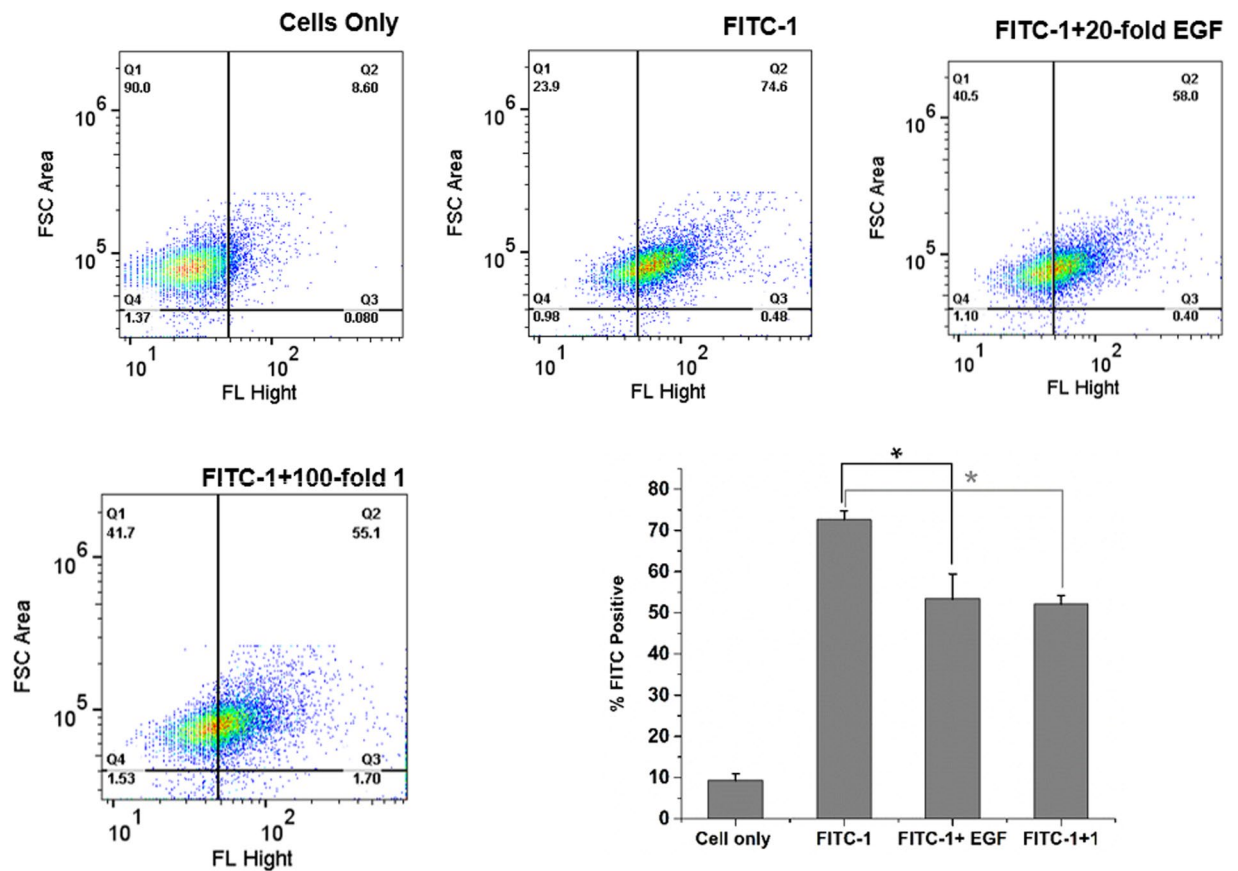
Peptide	Mean Fluorescence Intensity $\pm$ SD		
	MDA-MB-468	MDA-MB-231	MDA-MB-436
1	68 $\pm$ 45	42 $\pm$ 8	105 $\pm$ 11
4	23741 $\pm$ 3068	4791 $\pm$ 354	305 $\pm$ 58
22	8384 $\pm$ 735	1978 $\pm$ 522	41 $\pm$ 27
23	4661 $\pm$ 883	1631 $\pm$ 110	36 $\pm$ 14
26	6457 $\pm$ 557	1877 $\pm$ 67	16 $\pm$ 18
27	5248 $\pm$ 407	1683 $\pm$ 163	21 $\pm$ 17

**Figure 1.** Uptake of the FITC-labeled peptides by triple negative breast cancer cells MDA-MB-468, MDA-MB-231, and MDA-MB-436. Cells ( $10^5$  or  $10^6$ /well) were treated with peptides (1  $\mu$ M) and incubated for 30 min at 37  $^{\circ}$ C before analysis by FACS. The experiments were repeated twice and done in duplicates for each sample. Control group was treated with DMSO (5%). The table lists the mean fluorescence intensity (MFI) for the FITC positive cells for three independent experiments  $\pm$  SD.

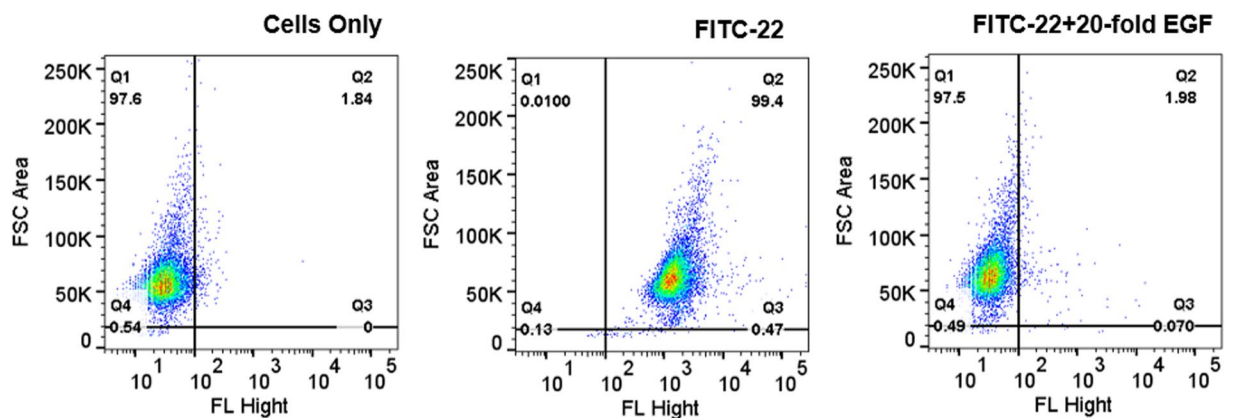
**Specific Uptake of Peptides by TNBC Cells.** The specific uptake of peptides **1** and **22** was evaluated using a competitive assay. First, MDA-MB-231 cells were incubated with peptide FITC-**1** in the presence and absence of excess EGF protein (20-fold). The cells showed about 17% decrease in fluorescence when incubated in the presence of EGF protein (Fig. 2). Next, the cells were incubated with peptide FITC-**1** with or without excess unlabeled peptide **1** (100-fold). A similar decrease in fluorescence (19.5% decrease) was observed when the cells were incubated with excess unlabeled peptide **1** suggesting multiple binding modes for peptide **1** on MDA-MB-231 cells. The competition experiment was also repeated with peptide **22** in the presence and absence of EGF protein. As shown in Fig. 3, the FITC positive cells reduced significantly from 99% to 2% (97% decrease) when the cells were incubated with peptide **22** in the presence of 20-fold excess EGF protein. The decrease in FITC positive cells was much higher for peptide **22** (97%) compared to peptide **1** (17%) suggesting that peptide **22** targets and binds EGFR than peptide **1**. These results support our conjecture that the newly designed peptide **22** targets and binds EGFR, likely at the same site as EGF protein, and the binding or uptake can be reduced when the receptor is blocked with excess EGF.

Further we incubated a fixed number of MDA-MB-231 cells ( $10^5$ ) with different concentrations of peptide **22** (0–6  $\mu$ M). The results showed that uptake was concentration independent in the range 1–6  $\mu$ M (Fig. S1) suggesting that half saturation is below 1  $\mu$ M. A direct interaction between the peptide and the isolated protein (EGFR) is required to confirm this specificity as well as the binding affinity. Previously Li *et al.* reported a  $K_d$  of 22 nM for peptide **1** binding to EGFR where the authors studied direct binding of the peptide to the pure EGFR protein<sup>8</sup>.

**Peptide 22 Structure.** As shown in Fig. 4a, the main difference in the chemical structures of peptides **1** and **22** is the presence of a negatively charged carboxylate in the side chain of glutamic acid of peptide **22**, whereas in peptide **1** this residue is uncharged (Gln). We next compared the sequence of peptide **22** with the EGF protein to understand the enhanced binding of peptide **22** to EGFR compared to peptide **1**. We found a region in EGF that may have structural similarity to peptide **22** (Fig. 4b). There is low sequence homology between the two sequences (peptide **22** and 12-residue region of EGF), however, based on the predicted secondary structure of peptide **22** and x-ray protein structure of EGF<sup>35</sup> it seems that these two have high structural similarity. Notably, this region of EGF (32–43) contains several residues that are involved in direct interaction with the EGFR based on the site-directed mutagenesis studies as well as the crystal structure of human EGF in complex with EGFR extracellular region<sup>36</sup>. Residues 32–43 of EGF loop to form a cyclic structure (Fig. 4c) and secondary structure prediction<sup>37,38</sup> of peptide **22** suggests that it consists of sheet and coil structures with a turn at P8 and E9 (Fig. 4d). Further, a homology based energy minimized structure for peptide **22** was obtained that shows structural similarity between residues 32–43 of EGF and peptide **22** (Fig. 4e). This structure displays that peptide **22** has a turn at E9 near the C-terminus and the peptide C-terminus coils back toward the N-terminus. These results suggest that a cyclic analogue of peptide **22** may give a high affinity EGFR binding peptide which should

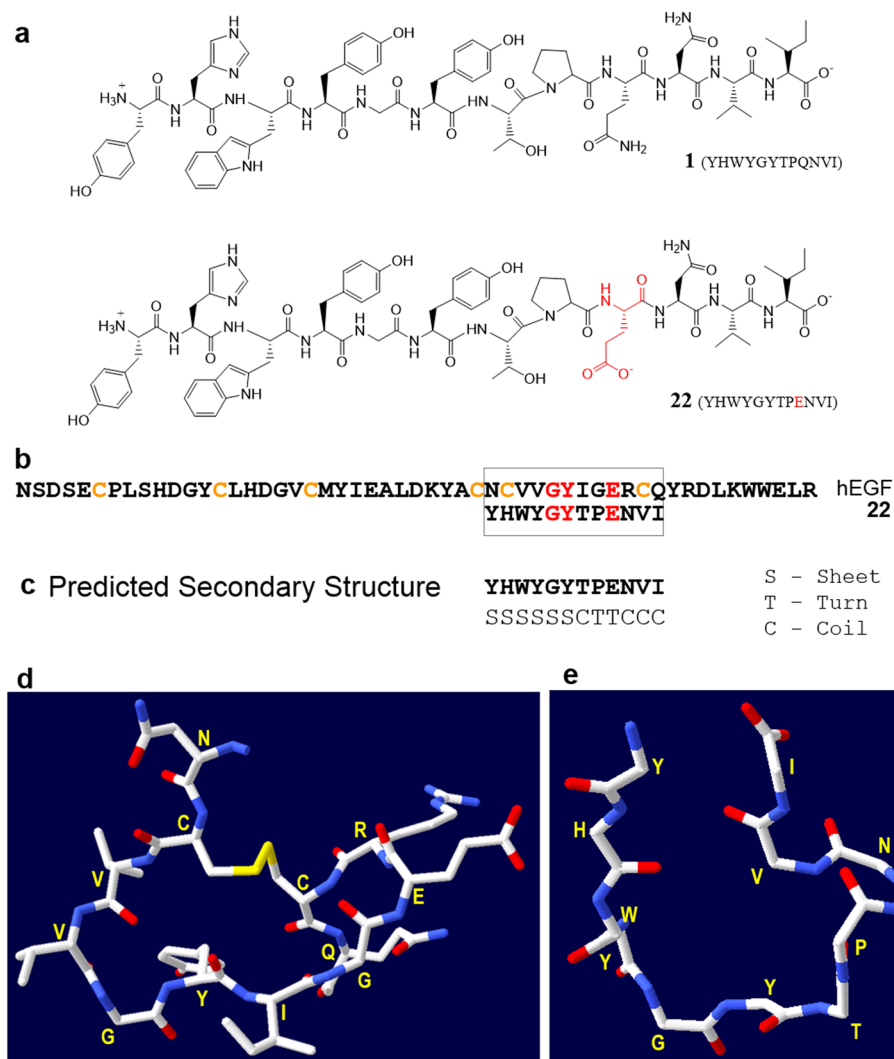


**Figure 2.** Competitive uptake of peptide 1 by FACS analysis. Fluorescence was measured for the MDA-MB-231 cells ( $10^5$  cells) alone or for the cells after incubation with peptide FITC-1 ( $1 \mu\text{M}$ ), peptide FITC-1 in the presence of 20-fold excess EGF protein ( $20 \mu\text{M}$ ) or peptide FITC-1 in the presence of 100-fold excess unlabeled peptide 1 ( $100 \mu\text{M}$ ). Also shown is the bar diagram for the percent of FITC positive cells in the absence and presence of excess free peptide 1 or EGF. Experiment was done in duplicates for each sample  $\pm$  SD. The asterisk (\*) denotes statistically significant difference ( $p < 0.05$ ). Each well of cells was read twice in the FACS instrument. 10,000 cells were read.



**Figure 3.** Competitive uptake of peptide 22 by FACS analysis. Fluorescence was measured for the MDA-MB-231 cells ( $10^5$  cells) alone or for the cells after incubation with peptide FITC-22 ( $1 \mu\text{M}$ ) or peptide FITC-22 in the presence of 20-fold excess EGF protein ( $20 \mu\text{M}$ ). Experiments were done in duplicates for each sample. 10,000 cells were read.

also have enhanced proteolytic stability. Such cyclic analogue of peptide 22 need to be tested. Peptide 22 shows short half-life ( $\sim 1$  hour) when incubated with human serum most likely due to proteolysis (data not shown). Previously we showed that a cyclic analogue (cyclic WXEAAYQkFL) of breast cancer targeting peptide 18-4



**Figure 4.** (a) Chemical structure of peptide **1** (GE11) and peptide analogue **22**. (b) Amino acid sequence of 53-mer human EGF (hEGF) and 12-mer **22** showing sequence homology for the G, Y and E residues (highlighted in red). Peptide **22** also shows structural similarity (based on secondary structure prediction of **22**) to this cyclic region of hEGF (residues 32–43). (c) Predicted secondary structure of peptide **22** using Chou and Fasman secondary structure prediction (CFSSP) server<sup>37,38</sup>. (d) 3D structure of a portion (residues 32–43) of hEGF derived from the crystal structure of full length hEGF (PDB code 1JL9)<sup>35</sup>. (e) Energy minimized structure of peptide **22** obtained using PEP-fold 2.0<sup>43</sup> and Swiss-Pdb Viewer 4.1<sup>42</sup>. Side chains are not shown here for clarity.

(WxEAA<sub>x</sub>YQrFL, where lower case x is D-norleucine or D-Nle) was stable toward proteolytic degradation as well as displayed higher uptake by the breast cancer cells compared to the linear peptide<sup>25</sup>.

### Concluding Remarks

Four analogues (**22**, **23**, **26**, and **27**) of peptide **1** (YHWYGYTPQNVI) with high uptake by TNBC cells that express high levels of EGFR are reported in the current investigation. Among these four new analogues, peptide **22** showed the highest uptake by the TNBC cells. For both MDA-MB-468 and MDA-MB-231 cells, a 123-fold and 47-fold increase, respectively, in uptake of soluble FITC-labeled peptide **22** was observed compared to the lead peptide **1**. Both these cell lines (MDA-MB-468 and MDA-MB-231) are reported to express high levels of EGFR. Mueller *et al.* used lysates of breast cancer cell lines to obtain proteins which were separated by SDS-PAGE and immunoblotted with anti-EGFR to show overexpression of EGFR by these TNBC cell lines<sup>14</sup>. The specific binding and uptake of peptide **22** by EGFR in TNBC cells is further confirmed by competition assay where the uptake of **22** (99%) was reduced to 2% when the cells were incubated in the presence of excess EGF protein.

The comparison of peptide **22** sequence with a portion of EGF protein suggests that there may be structural similarity further supporting peptide's specificity for EGFR. We are currently investigating an N- to C-terminal cyclic analogue of peptide **22**. Our preliminary results for the stability of peptides **1** and **22** in the presence of human serum show that these peptide have a short half-life (data not shown here). Our conjecture is that cyclization of peptide **22** will impart proteolytic stability while maintaining specificity for EGFR for its use in TNBC

drug targeting. Targeted delivery of doxorubicin for the treatment of ovarian cancer was recently reported using peptide GE11 conjugated reversibly crosslinked polymersomes carrying doxorubicin<sup>39</sup>. These peptide GE11 conjugated polymersomes showed 2.5-fold higher accumulation in the tumor compared to liposomal doxorubicin, caused minimal side effects at high doses, and led to increased survival rate in SKOV3-tumor bearing mice<sup>39</sup>. The new analogues of peptide 1 (or GE11) such as peptides 22, 23, 26, and 27 could serve as superior ligands for the targeted delivery of therapeutics for cancer treatment.

## Experimental Section

**Materials.** All solvents and chemicals including triisopropylsilane (TIPS), dichloromethane (DCM), dimethylformamide (DMF), N-methylmorpholine (NMM), diethyl ether, ethanol, acetonitrile, trifluoroacetic acid (TFA) were obtained from Sigma-Aldrich (now Millipore Sigma, St. Louis, MO, USA). Piperidine (20%) was purchased from Protein Technology (AZ, Tucson, USA). CyQUANT and FITC dyes were obtained from Invitrogen (Eugene, Oregon, USA). Derivatized cellulose membranes were from Intavis (AG, Germany). Milli-Q system was used for ultrapure water. Media RPMI-1640 and Leibovitz's L-15 were from American Type Culture Collection (ATCC, Manassas, VA, USA). DMEM, Step/Pen, Hank's Balanced Salt Solution (HBSS), 0.05% trypsin/EDTA and horse serum were obtained from Gibco (USA) and Hyclone (Pittsburgh, PA, USA). Hydrocortisone was from Acros Organics (New Jersey, USA). Human recombinant EGF was obtained from Corning (New York, USA).

Peptide array syntheses were done on automated spot synthesizer ResPep SL apparatus (Intavis AG, Germany). Synthesis of soluble peptides was performed on automated peptide synthesizer Tribute from Protein Technologies (AZ, Tucson, USA). HPLC system used was Prominence-i Shimadzu (Kyoto, Japan). ChemiDoc™ XRS + system from Bio-Rad (California, USA) was used to record fluorescence intensity of spots on membranes. Cell uptake studies were carried out on BD FACSVerser Flow Cytometer (BD Biosciences, California, USA). Cells were counted using hemocytometer Neubauer Chamber. Peptide concentration was determined using Shimadzu BioSpec-nano Micro-volume UV-Vis Spectrophotometer (Shimadzu, Kyoto, Japan).

**Cell Lines.** Cell lines were purchased from ATCC and were treated as recommended by ATCC for unpacking, handling, and culturing. MCF-10A and MCF-12A were cultured in DMEM/F12 media containing all components needed for growth media (Horse serum 5%, EGF 20 ng/mL, hydrocortisone 0.5 μg/mL, insulin 10 μg/mL, Pen/Strep 1%). The media used for these two cell lines are modified version of ATCC. For culturing MDA-MB-436, L15 media supplemented with 10 μg/mL insulin, 16 μg/mL glutathione and 10% FBS was used. MDA-MB-231 and MDA-MB-468 were cultured in DMEM containing 10% FBS and 1% Pen/Strep. Except for MDA-MB-436, all cells were cultivated at 37 °C with 5% CO<sub>2</sub>. MDA-MB-436 cells were kept at 37 °C with 100% air. The conditions for cell culture for each of the cell lines are described in Supplementary Information Table S1.

**Peptide Array Synthesis.** Thirty peptide sequences in duplicates were synthesized on functionalized cellulose membrane using fully automated ResPep SL peptide array synthesizer (Intavis, Germany) following previous procedure with some modifications (Fig. S2)<sup>40,41</sup>. Previously we used a semi-automatic AutoSpot (Intavis) synthesizer that required the fmoc deprotection with piperidine to be performed manually<sup>32</sup>. However, in the fully automated ResPep SL peptide array synthesizer all the steps for building the peptide chain are done by the robot. Briefly, the sequences of peptides were entered in robot software and 96 spot synthesis mode was selected for the library. The membrane was first immersed in DMF for an hour. It was then placed inside the holder of robot on a paper filter and was fixed with screws. The membrane was washed with ethanol few times and dried using vacuum to remove any bubbles below the membrane. Fmoc-peptide synthesis methodology with DIC and Oxyma pure as coupling reagent and racemization suppressor, respectively, was used. For each peptide spot, an 80 μL aliquot of activated amino acid (0.5 M) was delivered on the membrane surface using a robotic syringe. After each amino acid was double coupled, capping of unreacted amino acids was done using acetic anhydride (3%) and subsequently Fmoc-deprotection was achieved using 20% piperidine. After the synthesis was complete, the membrane was removed from the synthesizer platform and was manually dipped in deprotection solution containing approximately 50% TFA (15 mL), DCM (15 mL), TIPS (900 μL) and water (600 μL), and was kept for 3 h at room temperature<sup>32</sup>. The deprotection of the side-chain protecting groups was done using the same reaction conditions as originally developed. This procedure allowed complete removal of all side chain protections including the side chain Pbf of Arg<sup>32</sup>. Thereafter the membrane was washed extensively in several steps with DCM, DMF and ethanol, respectively. Finally it was air dried and kept in sealed bag at -20 °C until use.

**Peptide Array-Cell Binding Assay.** For the binding assay, the peptide array was incubated with the cells in sequential steps as described previously<sup>32</sup>. Briefly peptide array membrane was dipped in ethanol for 30 s to remove any precipitation on it, followed by dipping in sterile PBS (pH 7.4) for another 30 min. The cells (75 × 10<sup>3</sup> cells/mL) were seeded directly on the membrane in a sterile plate and incubated at 37 °C for 4 h. The unbound cells were washed out with sterile PBS (pH 7.4). Thereafter the membrane was placed in a sealed bag and transferred to -80 °C freezer for 2 h. Subsequently it was defrosted at room temperature and treated with CyQUANT solution. The CyQUANT concentration was optimized based on the manufacturer's protocol and preliminary experiments. Membrane was incubated at 37 °C for 30 min with the dye. After washing with sterile PBS and air drying using blow dryer, the membrane was scanned using Bio-rad chemidoc imager with image lab software. The settings were adjusted based on the excitation (465 nm) and emission (535 nm) wavelengths of CyQUANT dye. Epi Blue was selected for the light source (455–485 nm) and filter number 4 (blue filter, 518–546 nm) was employed. The fluorescence intensity of cancer cells read using chemidoc imager was normalized using the non-cancer cell fluorescence intensity. Eventually relative fluorescence was assessed to select peptides with high affinity to cancer cells and low or minimal binding to normal cells.

**Synthesis of FITC Labeled Peptides.** Six peptides, including peptide **1** (GE11), peptide **4** (p160) and analogues **22**, **23**, **26** and **27**, were selected for further studies. These peptides were synthesized following Fmoc solid phase peptide synthesis (Fmoc-SPPS) on preloaded Fmoc-isoleucine or leucine Wang resin (0.1 mmol scale) using automated peptide synthesizer (Tribute, Protein Technology) as reported previously. Preloaded Fmoc-isoleucine or leucine Wang resin (320 mg, 0.1 mmol) was added to the glass reaction vessel (RV) and the resin was allowed to swell in DMF under nitrogen with mechanical shaking for 30 min. All amino acids (3 equiv in 3 mL DMF) were coupled in sequence using HCTU (2.5 equiv) and NMM (1.2 equiv) for each coupling. Fmoc was removed by 20% piperidine in DMF.  $\beta$ -alanine was added at the N-terminal of each peptide sequence as a spacer before FITC coupling. FITC (0.3 mmol) in DMF (3 mL) with DIPEA (0.15 mmol) was added to the resin, and the mixture was incubated in the dark for 20 h. Small amount of the resin was used for a test cleavage using cleavage cocktail (1 mL) of trifluoroacetic acid (TFA)/triisopropylsilane/ ultra-pure water (95:2.5:2.5) for 2 hours. Cold diethyl ether (Et<sub>2</sub>O, 5 mL) was added to the filtered resin followed by centrifugation for 10 min in order to precipitate and collect the crude peptide. The successful conjugation was confirmed using MALDI-TOF and RP-HPLC of the test cleavage. Cleavage of FITC-peptides from the entire resin was done manually using same cleavage cocktail (10 mL) as described above. The peptides were precipitated using cold diethyl ether (20 mL). All peptides were characterized using MALDI-TOF mass spectrometry and RP-HPLC (Table S2 and Figs S3–S8). Purification was done using semi-preparative C18 Vydac column RP-HPLC.

**In vitro Cellular Uptake.** Fluorescence-activated cell-sorting (FACS) analysis was used to check the uptake of the FITC-peptides. The stock solution for the FITC-peptides (100  $\mu$ M) was prepared by dissolving the peptides in sterile water/10% DMSO. MDA-MB-231, MDA-MB-468 and MDA-MB-436 cells were seeded in 6-well plates with 10<sup>6</sup>/10<sup>5</sup> cells/well containing 3 mL media and the plates were incubated at 37 °C overnight. The following day, the media was replaced with fresh FBS free media (1 mL) containing FITC-labeled peptide (1  $\mu$ M), and further incubated at 37 °C for 30 min. Thereafter the media was removed and the cells were washed with cold PBS to remove any remaining peptide. Adequate trypsin was added to each well followed by 3–5 min incubation at 37 °C. Trypsin was immediately neutralized by adding media and the cell suspension was transferred into centrifuge tubes. The cells were separated and washed twice with FACS sheath fluid. The control cells were treated similarly but without peptide. Two wells were assigned for each peptide and each well sample was read twice in FACS machine. The experiment was repeated once. MFI represents the average of all wells for each peptide.

In addition, MDA-MB-231 cells (10<sup>5</sup>) were incubated with varying concentration of peptide FITC-**22** (0–6  $\mu$ M). The resulting samples were treated as above, and the experiment was repeated. The fluorescence intensity of the FITC peptide-cell complex was plotted against peptide concentration. Finally, competition experiment was done in the presence of EGF (20  $\mu$ M, 20x excess) or unlabeled peptide **1** (100  $\mu$ M, 100 times more) as competitors. Protein EGF or unlabeled peptide was added to the wells first, followed by addition of FITC-labeled peptide **1** or peptide **22** and the plate was incubated for 30 min. Control group was treated with media alone containing ~5% DMSO. Each group was done in duplicate and the experiment was repeated once.

**Structure Prediction Methods.** The Chou and Fasman secondary structure prediction (CFSSP) server was used to obtain predicted secondary structures for peptide **22**<sup>37,38</sup>. The three dimensional structure of a portion of EGF (12 residues, 32–43) was created using the PDB structure of EGF protein (PDB code 1JL9)<sup>35</sup> and Swiss-Pdb Viewer 4.1<sup>42</sup>. A PDB structure for peptide **22** was generated using the de novo peptide structure prediction software PEP-fold 2.0<sup>43</sup>. The structure was further refined using the “magic fit” module and energy minimization of Swiss-Pdb Viewer 4.1<sup>42</sup>.

## References

- Irvin, W. J. Jr. & Carey, L. A. What is triple-negative breast cancer? *Eur J Cancer* **44**, 2799–2805, <https://doi.org/10.1016/j.ejca.2008.09.034> (2008).
- Palma, G. *et al.* Triple negative breast cancer: looking for the missing link between biology and treatments. *Oncotarget* **6**, 26560–26574, <https://doi.org/10.18632/oncotarget.5306> (2015).
- Parise, C. & Caggiano, V. The influence of marital status and race/ethnicity on risk of mortality for triple negative breast cancer. *PLoS One* **13**, e0196134, <https://doi.org/10.1371/journal.pone.0196134> (2018).
- Turkman, Y. E., Sakibia Opong, A., Harris, L. N. & Knobf, M. T. Biologic, demographic, and social factors affecting triple negative breast cancer outcomes. *Clin J Oncol Nurs* **19**, 62–67, <https://doi.org/10.1188/15.CJON.62-67> (2015).
- Collignon, J., Lousberg, L., Schroeder, H. & Jerusalem, G. Triple-negative breast cancer: treatment challenges and solutions. *Breast Cancer (Dove Med Press)* **8**, 93–107, <https://doi.org/10.2147/BCTT.S69488> (2016).
- Dawood, S. Triple-negative breast cancer: epidemiology and management options. *Drugs* **70**, 2247–2258, <https://doi.org/10.2165/11538150-000000000-00000> (2010).
- Koenders, P. G. *et al.* Epidermal growth factor receptor-negative tumors are predominantly confined to the subgroup of estradiol receptor-positive human primary breast cancers. *Cancer Res* **51**, 4544–4548 (1991).
- Li, Z. *et al.* Identification and characterization of a novel peptide ligand of epidermal growth factor receptor for targeted delivery of therapeutics. *FASEB J* **19**, 1978–1985, <https://doi.org/10.1096/fj.05-4058com> (2005).
- Sainsbury, J. R., Farndon, J. R., Needham, G. K., Malcolm, A. J. & Harris, A. L. Epidermal-growth-factor receptor status as predictor of early recurrence of and death from breast cancer. *Lancet* **1**, 1398–1402 (1987).
- Toi, M., Osaki, A., Yamada, H. & Toge, T. Epidermal growth factor receptor expression as a prognostic indicator in breast cancer. *Eur J Cancer* **27**, 977–980 (1991).
- Costa, R. *et al.* Targeting Epidermal Growth Factor Receptor in triple negative breast cancer: New discoveries and practical insights for drug development. *Cancer Treat Rev* **53**, 111–119, <https://doi.org/10.1016/j.ctrv.2016.12.010> (2017).
- Changavi, A. A., Shashikala, A. & Ramji, A. S. Epidermal Growth Factor Receptor Expression in Triple Negative and Nontriple Negative Breast Carcinomas. *J Lab Physicians* **7**, 79–83, <https://doi.org/10.4103/0974-2727.163129> (2015).
- Fitzpatrick, S. L., Lachance, M. P. & Schultz, G. S. Characterization of Epidermal Growth-Factor Receptor and Action on Human-Breast Cancer-Cells in Culture. *Cancer Research* **44**, 3442–3447 (1984).
- Mueller, K. L., Yang, Z. Q., Haddad, R., Ethier, S. P. & Boerner, J. L. EGFR/Met association regulates EGFR TKI resistance in breast cancer. *J Mol Signal* **5**, 8, <https://doi.org/10.1186/1750-2187-5-8> (2010).



15. Aina, O. H., Sroka, T. C., Chen, M. L. & Lam, K. S. Therapeutic cancer targeting peptides. *Biopolymers* **66**, 184–199, <https://doi.org/10.1002/bip.10257> (2002).
16. Laakkonen, P., Zhang, L. & Ruoslahti, E. Peptide targeting of tumor lymph vessels. *Annals of the New York Academy of Sciences* **1131**, 37–43, <https://doi.org/10.1196/annals.1413.003> (2008).
17. Soudy, R. *et al.* Engineered Peptides for Applications in Cancer-Targeted Drug Delivery and Tumor Detection. *Mini Rev Med Chem* **17**, 1696–1712, <https://doi.org/10.2174/1389557516666160219121836> (2017).
18. Sugahara, K. N. *et al.* Coadministration of a tumor-penetrating peptide enhances the efficacy of cancer drugs. *Science* **328**, 1031–1035, <https://doi.org/10.1126/science.1183057> (2010).
19. Arap, W., Pasqualini, R. & Ruoslahti, E. Cancer treatment by targeted drug delivery to tumor vasculature in a mouse model. *Science* **279**, 377–380 (1998).
20. Curnis, F. *et al.* Enhancement of tumor necrosis factor alpha antitumor immunotherapeutic properties by targeted delivery to aminopeptidase N (CD13). *Nat Biotechnol* **18**, 1185–1190, <https://doi.org/10.1038/81183> (2000).
21. Soudy, R., Ahmed, S. & Kaur, K. NGR peptide ligands for targeting CD13/APN identified through peptide array screening resemble fibronectin sequences. *ACS Comb Sci* **14**, 590–599, <https://doi.org/10.1021/co300055s> (2012).
22. Askoxylakis, V. *et al.* Preclinical evaluation of the breast cancer cell-binding peptide, p160. *Clin Cancer Res* **11**, 6705–6712, <https://doi.org/10.1158/1078-0432.CCR-05-0432> (2005).
23. Soudy, R., Etayash, H., Bahadorani, K., Lavasanifar, A. & Kaur, K. Breast Cancer Targeting Peptide Binds Keratin 1: A New Molecular Marker for Targeted Drug Delivery to Breast Cancer. *Mol Pharm* **14**, 593–604, <https://doi.org/10.1021/acs.molpharmaceut.6b00652> (2017).
24. Zhang, J., Spring, H. & Schwab, M. Neuroblastoma tumor cell-binding peptides identified through random peptide phage display. *Cancer Lett* **171**, 153–164 (2001).
25. Raghuvanshi, Y. *et al.* Proteolytically Stable Cyclic Decapeptide for Breast Cancer Cell Targeting. *J Med Chem* **60**, 4893–4903, <https://doi.org/10.1021/acs.jmedchem.7b00163> (2017).
26. Zhang, M. Z. *et al.* High transfection efficiency of quantum dot-antisense oligonucleotide nanoparticles in cancer cells through dual-receptor synergistic targeting. *Nanotechnology* **25**, 255102, <https://doi.org/10.1088/0957-4484/25/25/255102> (2014).
27. Chen, J. *et al.* EGFR and CD44 Dual-Targeted Multifunctional Hyaluronic Acid Nanogels Boost Protein Delivery to Ovarian and Breast Cancers *In Vitro* and *In Vivo*. *ACS Appl Mater Interfaces*, <https://doi.org/10.1021/acsami.7b06879> (2017).
28. Uusi-Kerttula, H. *et al.* Incorporation of Peptides Targeting EGFR and FGFR1 into the Adenoviral Fiber Knob Domain and Their Evaluation as Targeted Cancer Therapies. *Human Gene Therapy* **26**, 320–329, <https://doi.org/10.1089/hum.2015.015> (2015).
29. Rahmanian, N. *et al.* 99mTc-radiolabeled GE11-modified peptide for ovarian tumor targeting. *Daru* **25**, 13, <https://doi.org/10.1186/s40199-017-0179-8> (2017).
30. Hu, D. *et al.* GE11 peptide modified and reduction-responsive hyaluronic acid-based nanoparticles induced higher efficacy of doxorubicin for breast carcinoma therapy. *Int J Nanomedicine* **11**, 5125–5147, <https://doi.org/10.2147/IJN.S113469> (2016).
31. Fan, M. *et al.* Design and biological activity of epidermal growth factor receptor-targeted peptide doxorubicin conjugate. *Biomed Pharmacother* **70**, 268–273, <https://doi.org/10.1016/j.biopha.2015.01.027> (2015).
32. Ahmed, S., Mathews, A. S., Byeon, N., Lavasanifar, A. & Kaur, K. Peptide arrays for screening cancer specific peptides. *Anal Chem* **82**, 7533–7541, <https://doi.org/10.1021/ac1003085> (2010).
33. Soudy, R., Gill, A., Sprules, T., Lavasanifar, A. & Kaur, K. Proteolytically stable cancer targeting peptides with high affinity for breast cancer cells. *Journal of medicinal chemistry* **54**, 7523–7534, <https://doi.org/10.1021/jm200750x> (2011).
34. Subik, K. *et al.* The Expression Patterns of ER, PR, HER2, CK5/6, EGFR, Ki-67 and AR by Immunohistochemical Analysis in Breast Cancer Cell Lines. *Breast Cancer (Auckl)* **4**, 35–41 (2010).
35. Lu, H. S. *et al.* Crystal structure of human epidermal growth factor and its dimerization. *J Biol Chem* **276**, 34913–34917, <https://doi.org/10.1074/jbc.M102874200> (2001).
36. Ogiso, H. *et al.* Crystal structure of the complex of human epidermal growth factor and receptor extracellular domains. *Cell* **110**, 775–787 (2002).
37. Ashok Kumar, T. CFSSP: Chou and Fasman Secondary Structure Prediction server. *Wide Spectrum* **1**, 15–19 (2013).
38. Chou, P. Y. & Fasman, G. D. Prediction of protein conformation. *Biochemistry* **13**, 222–245 (1974).
39. Zou, Y., Xia, Y., Meng, F., Zhang, J. & Zhong, Z. GE11-Directed Functional Polymersomal Doxorubicin as an Advanced Alternative to Clinical Liposomal Formulation for Ovarian Cancer Treatment. *Mol Pharm* **15**, 3664–3671, <https://doi.org/10.1021/acs.molpharmaceut.8b00024> (2018).
40. Frank, R. The SPOT-synthesis technique. Synthetic peptide arrays on membrane supports—principles and applications. *J Immunol Methods* **267**, 13–26 (2002).
41. Azmi, S., Jiang, K., Stiles, M., Thundat, T. & Kaur, K. Detection of Listeria monocytogenes with Short Peptide Fragments from Class IIa Bacteriocins as Recognition Elements. *ACS Combinatorial Science* **17**, 156–163, <https://doi.org/10.1021/co500079k> (2015).
42. Guex, N. & Peitsch, M. C. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* **18**, 2714–2723, <https://doi.org/10.1002/elps.1150181505> (1997).
43. Shen, Y., Maupetit, J., Derreumaux, P. & Tuffery, P. Improved PEP-FOLD Approach for Peptide and Mini-protein Structure Prediction. *J Chem Theory Comput* **10**, 4745–4758, <https://doi.org/10.1021/ct500592m> (2014).

## Acknowledgements

Research reported in this publication was supported by the National Cancer Institute of the National Institutes of Health under Award Number R15CA208656. We thank the core labs at the Chapman University School of Pharmacy for access to all instrumentation. We thank Azam Saghaeidehkordi and Elmira Ziaei for help with the minor revision of the manuscript.

## Author Contributions

K.K. conceived and designed the experiments. H.A. performed most of the experiments, E.A. performed the peptide human serum stability experiment, and K.K. performed the structural studies and analysis thereof. K.K. and H.A. analyzed the results and wrote the manuscript. K.K. did all the work toward the major revision of the manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-019-38574-y>.

**Competing Interests:** The authors declare no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019