

## Rational Identification of a Colorectal Cancer Targeting Peptide through Phage Display

Ferreira, Débora; Silva, Ana P.; Nobrega, Franklin L.; Martins, Ivone M.; Barbosa-Matos, Catarina; Granja, Sara; Martins, Sandra F.; Baltazar, Fátima; Rodrigues, Ligia R.

**DOI**

[10.1038/s41598-019-40562-1](https://doi.org/10.1038/s41598-019-40562-1)

**Publication date**

2019

**Document Version**

Final published version

**Published in**

Scientific Reports

**Citation (APA)**

Ferreira, D., Silva, A. P., Nobrega, F. L., Martins, I. M., Barbosa-Matos, C., Granja, S., Martins, S. F., Baltazar, F., & Rodrigues, L. R. (2019). Rational Identification of a Colorectal Cancer Targeting Peptide through Phage Display. *Scientific Reports*, 9(1), Article 3958. <https://doi.org/10.1038/s41598-019-40562-1>

**Important note**

To cite this publication, please use the final published version (if applicable).  
Please check the document version above.

**Copyright**

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

**Takedown policy**

Please contact us and provide details if you believe this document breaches copyrights.  
We will remove access to the work immediately and investigate your claim.

# SCIENTIFIC REPORTS



OPEN

## Rational Identification of a Colorectal Cancer Targeting Peptide through Phage Display

Débora Ferreira<sup>1,2</sup>, Ana P. Silva<sup>1</sup>, Franklin L. Nobrega<sup>1,3</sup>, Ivone M. Martins<sup>1</sup>, Catarina Barbosa-Matos<sup>4,5</sup>, Sara Granja<sup>4,5</sup>, Sandra F. Martins<sup>4,5,6</sup>, Fátima Baltazar<sup>4,5</sup> & Ligia R. Rodrigues<sup>1,2</sup>

Colorectal cancer is frequently diagnosed at an advanced stage due to the absence of early clinical indicators. Hence, the identification of new targeting molecules is crucial for an early detection and development of targeted therapies. This study aimed to identify and characterize novel peptides specific for the colorectal cancer cell line RKO using a phage-displayed peptide library. After four rounds of selection plus a negative step with normal colorectal cells, CCD-841-CoN, there was an obvious phage enrichment that specifically bound to RKO cells. Cell-based enzyme-linked immunosorbent assay (ELISA) was performed to assess the most specific peptides leading to the selection of the peptide sequence CPKSNNGVC. Through fluorescence microscopy and cytometry, the synthetic peptide RKOpep was shown to specifically bind to RKO cells, as well as to other human colorectal cancer cells including Caco-2, HCT 116 and HCT-15, but not to the normal non-cancer cells. Moreover, it was shown that RKOpep specifically targeted human colorectal cancer cell tissues. A bioinformatics analysis suggested that the RKOpep targets the monocarboxylate transporter 1, which has been implicated in colorectal cancer progression and prognosis, proven through gene knockdown approaches and shown by immunocytochemistry co-localization studies. The peptide herein identified can be a potential candidate for targeted therapies for colorectal cancer.

Colorectal cancer (CRC) is the third most commonly diagnosed cancer worldwide and the second leading cause of cancer-related deaths<sup>1</sup>. The initiation and progression of benign adenoma to malignant adenocarcinoma may be driven by the accumulation of several gene mutations and epigenetic modifications<sup>2</sup>. Early stage screening of CRC can potentially reduce both the incidence and mortality from this type of cancer. However, due to limitations of the current screening modalities in CRC (colonoscopy, biopsy and blood tests), several efforts are being conducted to discover new biomarkers that could be used as alternative screening tools for early diagnosis. Amongst these, peptide ligands that specifically recognize cell surface receptors are particularly promising and are being extensively used in cancer research. Peptides have become an attractive alternative, as they are easy to synthesize in large amounts and their small size improves tissue penetration, with less nonspecific uptake by the reticuloendothelial system<sup>3</sup>. Moreover, they can be chemically modified to alter affinity, charge, hydrophobicity, stability, and solubility and have been used to functionalize different nanosystems for improved and targeted therapy<sup>4</sup>.

Peptides can be selected in a relatively cost-effective manner using phage display<sup>5-9</sup>. This powerful technology was first introduced in 1985<sup>10</sup> and has been modified to a rapid high-throughput one step method - Biopanning and Rapid Analysis of Selective Interactive Ligands (BRASIL)<sup>11</sup>, which has enabled the construction of a large number of phage peptide libraries, with a wide range of applications<sup>12-15</sup>. The phage display screening methodology has been exploited and applied in cancer progression and used in selection processes, not only on solid primary tumors, but also on tumor vasculature, metabolism, cell signaling targets and metastasis<sup>7,16,17</sup>. Furthermore, bioinformatics tools and web servers have proven useful to validate and characterize these novel ligands<sup>18</sup>.

<sup>1</sup>Centre of Biological Engineering, University of Minho (CEB), Campus de Gualtar, 4710-057, Braga, Portugal.

<sup>2</sup>MIT-Portugal Program, Lisbon, Portugal. <sup>3</sup>Present address: Department of Bionanoscience, Kavli Institute of Nanoscience, Delft University of Technology, Van der Maasweg 9, 2629 HZ, Delft, Netherlands. <sup>4</sup>Life and Health Sciences Research Institute (ICVS), School of Health Sciences, University of Minho, Campus de Gualtar, 4710-057, Braga, Portugal. <sup>5</sup>ICVS/3B's - PT Government Associate Laboratory, Braga/Guimarães, Portugal. <sup>6</sup>Surgery Department, Coloproctology Unit, Braga Hospital, Braga, Portugal. Correspondence and requests for materials should be addressed to L.R.R. (email: [lrmr@deb.uminho.pt](mailto:lrmr@deb.uminho.pt))

Rounds	Input phages (PFU/mL)	Output phages (PFU/mL)	Ratio (output/input)
1	$1.0 \times 10^{11}$	$2.5 \times 10^8$	$2.5 \times 10^{-3}$
2	$2.5 \times 10^8$	$6.1 \times 10^5$	$2.44 \times 10^{-3}$
3	$6.1 \times 10^5$	$1.0 \times 10^4$	$1.64 \times 10^{-2}$
4	$1.1 \times 10^{10}$	$6 \times 10^8$	$5.5 \times 10^{-2}$
5 (Negative Selection)	$1.0 \times 10^{10}$	$1.1 \times 10^9$	$1.1 \times 10^{-1}$

**Table 1.** Enrichment of RKO cell-bound phages for each round of selection.

Sequence ID	Sequence	Round
RKO-R3-1	C--IAVPSNL-C	3
RKO-R3-2	C--KTPNGHLC	3
RKO-R3-3	C--KXPKGHLC	3
RKO-R4-1	CRPTYSPS---C	4
RKO-R4-2	C-KIHSSET--C	4
RKO-R4-3	C--PKSNNGV-C	4
RKO-R4-4	C--IGNSNTL-C	4
RKO-R5-1	C--PKSNNGV-C	5
RKO-R5-2	C--KTPNGHLC	5
RKO-R5-3	CQSISTAH---C	5
RKO-R5-4	C-NDDVPNK--C	5
RKO-R5-5	C--EIPGKVVC	5
RKO-R5-6	C--LRTPANH-C	5

**Table 2.** Multiple sequence alignment of peptides identified as candidates in rounds 3, 4 and 5 of biopanning generated by Clustal Omega.

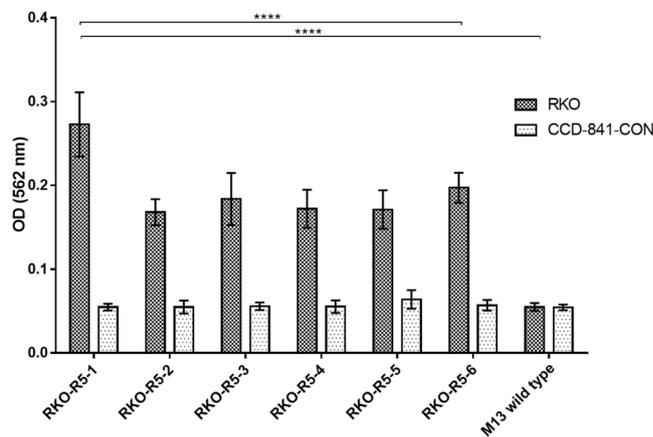
Herein we used phage display to identify a peptide, RKOpep, that specifically binds to the cell surface of the human CRC cell lines RKO, Caco-2, HCT 116 and HCT-15, as well as to colorectal cancer tissues. Monocarboxylate transporter 1 (MCT1) was suggested as a possible target based on a bioinformatics analysis and it was further confirmed by gene downregulation approaches and immunocytochemistry co-localization studies. Our results propose a novel targeting system for CRC diagnosis and/or treatment.

## Results

**Specific enrichment of RKO-binding phages.** A total of four rounds of selection with RKO cells were performed through *in vitro* biopanning, followed by a negative selection step against normal colon CCD-841-CoN cell line. In each round, the phages that specifically bound to target cells were recovered and used for the next round of selection. In the three initial rounds of selection, the obtained phage pool was not amplified between rounds. However, due to loss of phage concentration, the phage particles obtained in the last two biopanning rounds were amplified using an engineered *E. coli* JM109<sup>+</sup> strain to minimize the presence of biased sequences<sup>19</sup>. The phage enrichment rate (output/input phage concentration) was gradually increased during the selection rounds, reaching a 45-fold increase at the final round (Table 1).

**DNA sequencing.** After five rounds of biopanning, RKO cell-bound phages were grown to obtain single plaques from which individual phage clones were amplified and genomic DNA extracted. Ten individual clones from rounds 3, 4 and 5 were sequenced and the amino acid sequences deduced and aligned using Clustal Omega (Table 2). Most of these clones displayed a wild type sequence, while the others had distinct peptide hits in which the sequences CKTPNGHLC (RKO-R5-2) and CPKSNNGVC (RKO-R5-1) appeared twice (Table 2). Similar sequences were aligned resulting in the identification of small motif sequences. For instance, lysine (Lys) appeared in the fourth position in 5 of the 13 peptide sequences. Moreover, proline (Pro) appeared eight times at the sixth position, asparagine (Asn) was present six times at the seventh site, and glycine (Gly) appeared in 5 sequences at the eighth position, suggesting a selective pressure for binding to a cell surface epitope.

**Binding assays of phage clones towards RKO cells by cell-based ELISA.** A cell-based ELISA assay was used to evaluate the binding capacity to RKO cells of different phages selected after the last round of biopanning (Fig. 1). This experiment excluded false positive colonies, with binding affinity similar to that of the wild-type phage ( $p < 0.0001$ ), and assessed binding selectivity by comparing the binding abilities of each phage clone to RKO and CCD-841-CoN cells. No significant binding of the six phage clones to the control cells CCD-841-CoN was found, suggesting the selectivity of all clones to the RKO cells. Of these, the clone RKO-R5-1 (CPKSNNGVC) showed the highest binding affinity ( $p < 0.0001$ ) being selected for further characterization.



**Figure 1.** Assessment by cell-ELISA of the binding selectivity to RKO and CCD-841-CoN cells of six phage clones from the last round of biopanning. M13 wild type phage without any displayed peptide was used as negative control. All data are expressed as the mean  $\pm$  SD of three independent experiments. Two-way ANOVA indicates statistically significant differences within the group assessed by Tukey post-test and denoted as follows: \*\*\*\* $p < 0.0001$ .

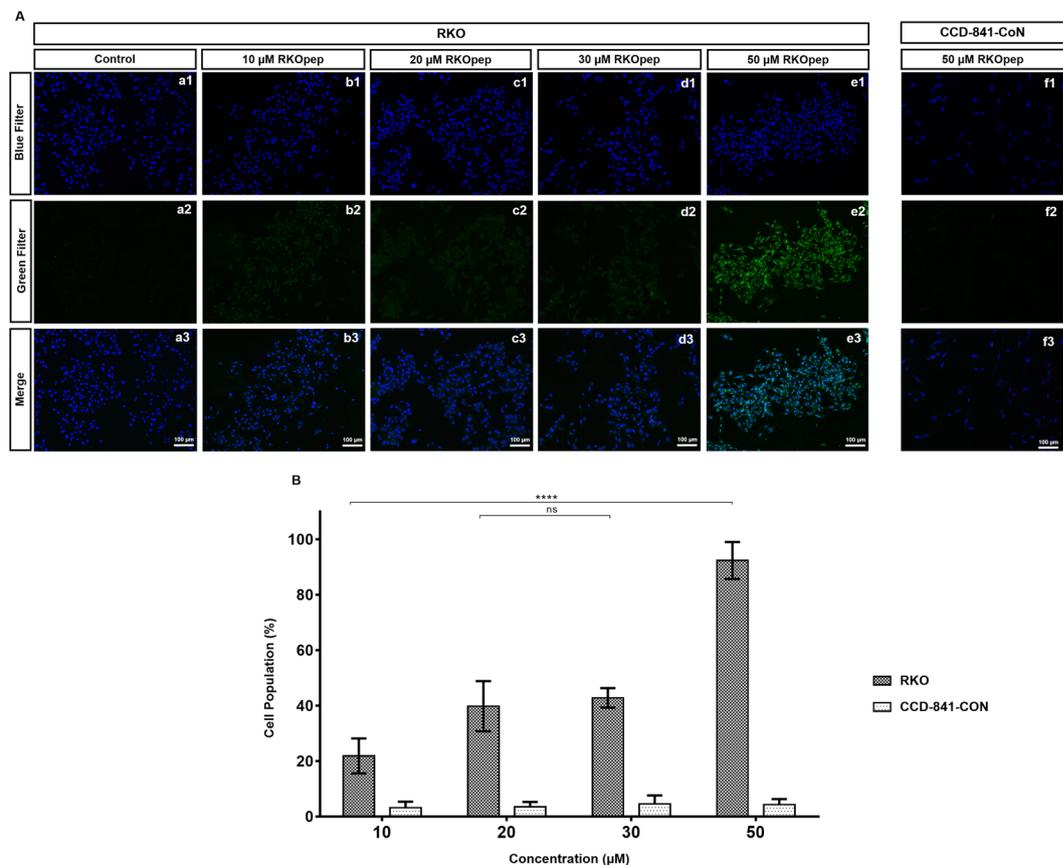
**Peptide analysis.** To accurately eliminate the existence of target unrelated peptides (TUPs), false positives and already existing mimotopes, the peptide sequence CPKSNNGVC was analyzed using some tools from Scanner and Reporter of Target-Unrelated Peptides (SAROTUP) that is the most extensive bioinformatics source for assessing peptides derived from phage display selections (Supplementary Table S1). No TUP hit sequences similar to the peptide described in this work were identified being more likely to be a true binder with high specificity to the target. Moreover, the identified peptide was predicted not to bind to contaminants or other components of the screening system (selection-related TUP) and also not to be a clone with growth advantage (propagation-related TUP)

**In vitro targeting of RKOpep to human colorectal cancer.** To study whether the free peptide (non-phage-displayed) maintained the binding ability and specificity shown in cell-based ELISA assays, the peptide CPKSNNGVC was synthesized with a FAM label (FAM-RKOpep) at the N-terminus. RKO and CCD-841-CoN cells were incubated with several working concentrations (10  $\mu$ M, 20  $\mu$ M, 30  $\mu$ M and 50  $\mu$ M) of FAM-labelled peptide and the results were evaluated under fluorescence microscopy and cytometry (Fig. 2A,B). The microscopy and cytometry results are in good agreement, i.e. fluorescence intensity increased with increasing concentrations of FAM-RKOpep in RKO cells in comparison with the control cells. For the higher FAM-RKOpep concentration tested (50  $\mu$ M), about 90% of the overall RKO cell population was bound by the RKOpep. It is also shown that for CCD-841-CoN, almost no positive signal was detected independently of the peptide concentration, thus confirming the specificity of the peptide RKOpep. In addition, to determine if the FAM-RKOpep specificity was limited to the human CRC cell line RKO, the peptide was incubated with other CRC cells (Fig. 3). A similar binding ability was observed, however, the binding affinity of the peptide at 50  $\mu$ M was significantly lower for Caco-2 ( $p < 0.001$ ), HCT 116 ( $p < 0.0001$ ) and HCT-15 ( $p < 0.0001$ ) cells when compared to the signal obtained for the RKO cell line. About  $56.69 \pm 6.8\%$ ,  $43.29 \pm 10.78\%$  and  $42.13 \pm 4.08\%$  of the overall Caco-2, HCT 116 and HCT-15 cell population was bound to the RKOpep, respectively. These findings proved the successful binding of the peptide RKOpep to different human CRC cells, although with variable binding abilities.

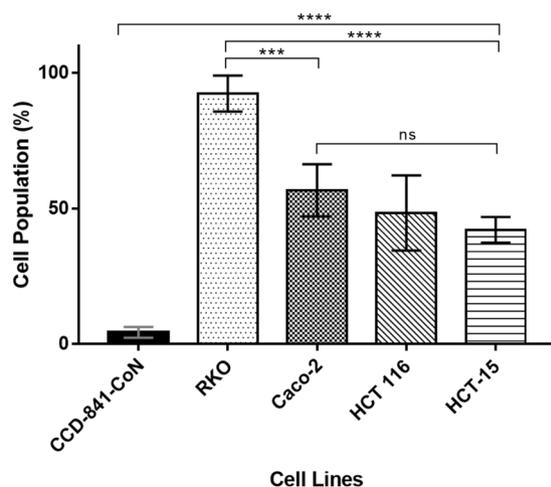
The tumor affinity of FAM-labelled RKOpep was also assessed using human colorectal cancer tissues through immunofluorescence staining (Fig. 4). Visible FAM signal, represented in green, resulting from the incubation of FAM-RKOpep was detected in tumor tissues ( $n = 3$ ). Using the same microscopy fluorescence settings after acquiring the images, a weak fluorescence for normal-adjacent tissues was observed, thus validating the specificity and the binding capacity of the peptide herein selected by the BRASIL method.

**Bioinformatics identification of the potential target for RKOpep.** A structural bioinformatics approach was used to identify the potential targets of RKOpep. A complete overview of all cell surface proteins of five CRC cell lines, including RKO, Caco-2 and HCT 116, was provided by de Wit and collaborators<sup>20</sup>. Using a high-resolution shotgun proteomics analysis, a total of 2609 proteins were identified in the cell surface fractions. By combining additional selection criteria including the presence in three of the cell lines in study, overexpression in carcinomas when compared to adenomas and subcellular location at the plasma membrane, about 59 candidate proteins for the RKOpep were identified.

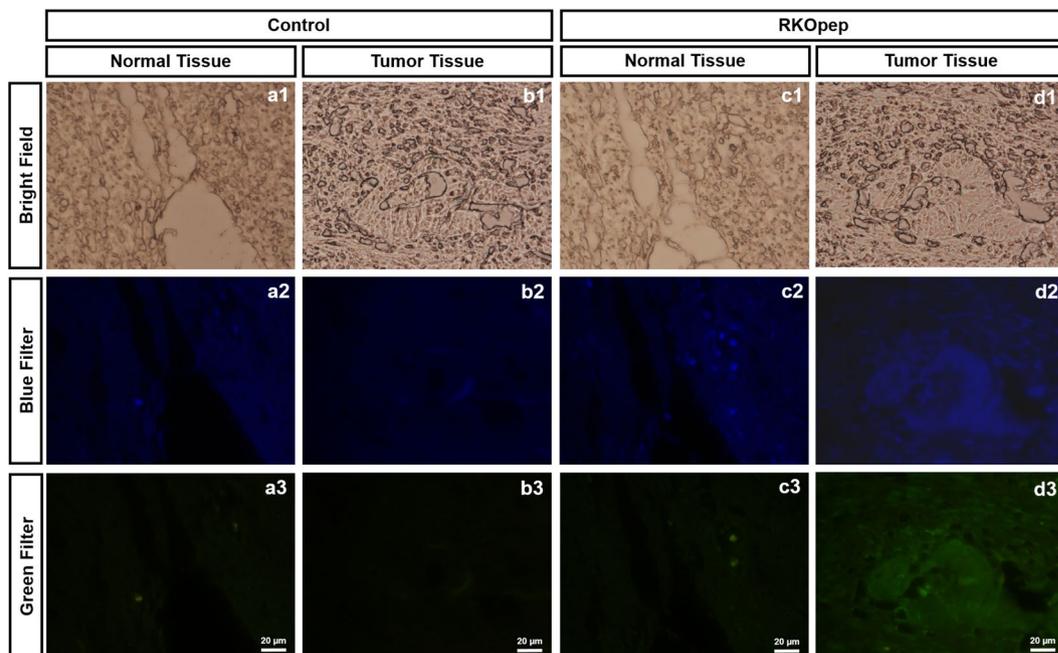
The peptide RKOpep was analyzed using protein-protein Basic Local Alignment Search Tool (BLAST) algorithm for homology to known colorectal cancer-related proteins. According to BLAST results, RKOpep potentially recognizes the membrane protein MCT1 that is one of the 59 candidate biomarkers that the proteomic analysis described by de Wit and coworkers<sup>20</sup> retrieved after applying selection criteria. The human protein target MCT1 was found with an overall identity of 100%, 0% gaps and with the highest E-value (Supplementary Table S2). In order to explore possible modes of binding of RKOpep to MCT1, computational docking was



**Figure 2.** Binding selectivity of FAM-RKOpep towards human colorectal cell line RKO and normal colon cell line CCD-841-CoN demonstrated by fluorescence microscopy and cytometry. **(A)** Microscopy results expressing the binding of increasing concentrations (10  $\mu\text{M}$ –50  $\mu\text{M}$ ) of FAM-RKOpep to RKO (b–e) and the highest concentration (50  $\mu\text{M}$ ) to CCD-841-CoN (f). The control corresponds to cells only stained with DAPI (a). (1) blue, nuclei stained with DAPI, (2) green, FAM-RKOpep and (3) overlapping of all filters (1 and 2). Scale bars represent 100  $\mu\text{m}$ . **(B)** Flow cytometry results expressing the binding of increasing concentrations (10  $\mu\text{M}$ –50  $\mu\text{M}$ ) of FAM-RKOpep to RKO and CCD-841-CoN. All data are expressed as the mean  $\pm$  SD of three independent experiments. Two-way ANOVA indicates statistically significant differences within the group assessed by Tukey post-test and denoted as follows: \*\*\*\* $p < 0.0001$  and ns  $p > 0.05$ .



**Figure 3.** Binding ability of FAM-RKOpep towards human colorectal cell lines demonstrated by cytometry. CCD-841-CoN, RKO, Caco-2, HCT 116 and HCT-15 cells were incubated with the highest concentration (50  $\mu\text{M}$ ) of FAM-RKOpep. All data are expressed as the mean  $\pm$  SD of three independent experiments. Two-way ANOVA indicates statistically significant differences within the group assessed by Tukey post-test and denoted as follows: \*\*\*\* $p < 0.0001$ , \*\*\* $p < 0.001$  and ns  $p > 0.05$ .



**Figure 4.** *In vitro* binding affinity of FAM-RKOpep to human colorectal cancer tissues through immunofluorescence staining. Control stands for experiments run without FAM-labelled RKOpep for normal (a) and cancer (b) tissues; RKOpep stands for experiments run with RKOpep for normal (c) and cancer (d) tissues. (1) bright field, (2) blue, nuclei stained with DAPI and (3) green, FAM-RKOpep. Scale bars represent 20  $\mu\text{m}$ .

performed using the flexible and fully blind protein-peptide docking software CABS-dock<sup>21</sup>. This software uses randomly generated peptide conformations, randomly orients these peptides over the protein surface and refines them using Monte Carlo dynamics<sup>22,23</sup>. The three-dimensional structure of MCT1 was predicted with 100% of confidence by using PHYRE2 software<sup>24</sup>. The best three-dimensional docking model, ranked according to trajectory characteristics, is shown in Fig. 5A. The 2D details of the intermolecular interactions between docked RKOpep peptide and MCT1 using LigPlot+ are shown in Fig. 5B.

In order to correlate the binding of RKOpep to the CRC cell lines with the expression of MCT1, its level of expression in the various cell lines under study, including RKO, Caco-2, HCT 116, HCT-15 and CCD-841-CoN, was evaluated through western blot (Fig. 6). Higher levels of MCT1 expression were verified for the CRC cell lines comparing to the normal one (control), being the MCT1 level of expression about 6–10 times superior than the one for CCD-841-CoN cells, according to band intensity calculations using ImageJ.

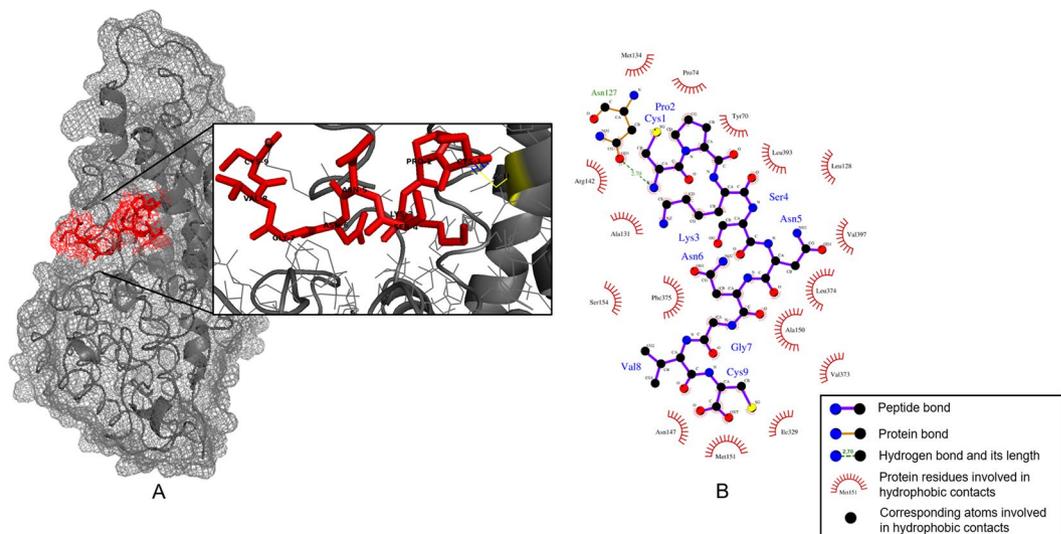
**Validation of MCT1 as a target of RKOpep.** To support and validate the bioinformatics analysis, gene knockdown approaches and immunocytochemistry studies were performed. As observed in Fig. 7, a clear reduction of the fluorescence signal (about 75%) on RKO cells was observed after MCT1 silencing when compared with the control ( $p < 0.0001$ ). To understand the selectivity of RKOpep to MCT1, silencing of another MCT isoform was also performed, namely monocarboxylate transporter 4 (MCT4). A fluorescence signal decrease of approximately 13% was seen after the MCT4 silencing ( $p < 0.001$ ). The scrambled peptide FAM-SCRpep exhibited low binding to the RKO cells and almost the same fluorescence signal, regardless of the conditions tested.

To demonstrate the RKOpep-MCT1 binding, immunocytochemistry experiments were also conducted. As observed in Fig. 8, the FAM-labelled RKOpep (represented in green) and anti-MCT1 antibody conjugated with a secondary antibody coupled to Alexa Fluor Plus 680 (represented in red) have fluorescence signal on RKO cells. The fluorescence signals overlapped after merging as demonstrated by the yellow/orange color obtained and seemed to co-localize on cell surface, supporting their specific interaction. Control samples were characterized by a faint or no staining, thus strengthening the validation of specific binding of the selected peptide to the MCT1 membrane protein.

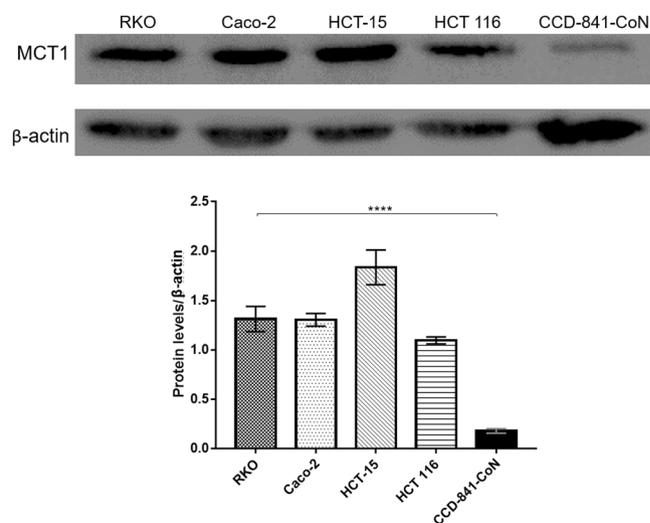
## Discussion

CRC is one of the worldwide leading causes of cancer-related morbidity and mortality<sup>25</sup>. Identifying new ligands that specifically target CRC may unravel novel perspectives to develop unique targeted therapies. Human CRC cell lines are useful preclinical model systems as they closely resemble primary tumors<sup>26</sup>. Herein we report the use of a modified phage display methodology to select peptides specific for the CRC cell line RKO. To our knowledge, this study is the first reported in the literature regarding RKO targeting, strongly supporting the need to find new targeting systems for CRC.

RKO-specific peptides were identified by BRASIL using four biopanning rounds of selection against RKO cells and a final negative selection round towards a normal colon cell line, to exclude all peptides that bind non-specifically to cancer cells. To accomplish this, a commercial combinatorial library of seven random amino



**Figure 5.** Binding mode of RKOpеп with MCT1. **(A)** Most favorable pose of 3D interaction between the RKOpеп and the MCT1 (by Pymol). This model was generated by CABS-dock, based on the PHYRE2 predicted structure of MCT1. The RKOpеп is shown as a red stick model. MCT1 is represented in grey with key residues shown in yellow. Blue dashed lines represent the hydrogen bonds. **(B)** Two-dimensional interaction between the RKOpеп and the MCT1 (by LigPlot). Residues with purple bonds in blue labels are from the RKOpеп; residues with brown bonds in red labels are from MCT1. The green dashed line represents a potential hydrogen bond and its length. Hydrophobic interactions are represented by arcs and radiating spokes.

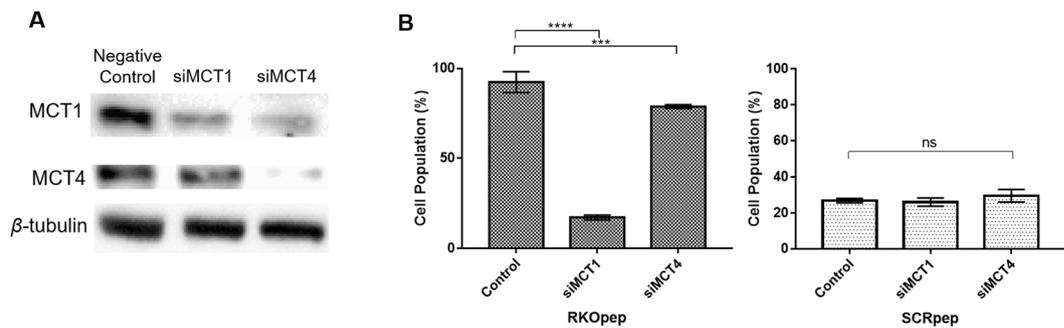


**Figure 6.** Expression of the MCT1 in colorectal cell lines. Detection of MCT1 (47 kDa) by western blot, in cell lysates of RKO, Caco-2, HCT-15, HCT 116 and CCD-841-CoN.  $\beta$ -actin (42 kDa) was used to confirm equal protein loading. Band intensity was measured using ImageJ software program. Similar results were obtained from three independent experiments and a representative blot is shown.

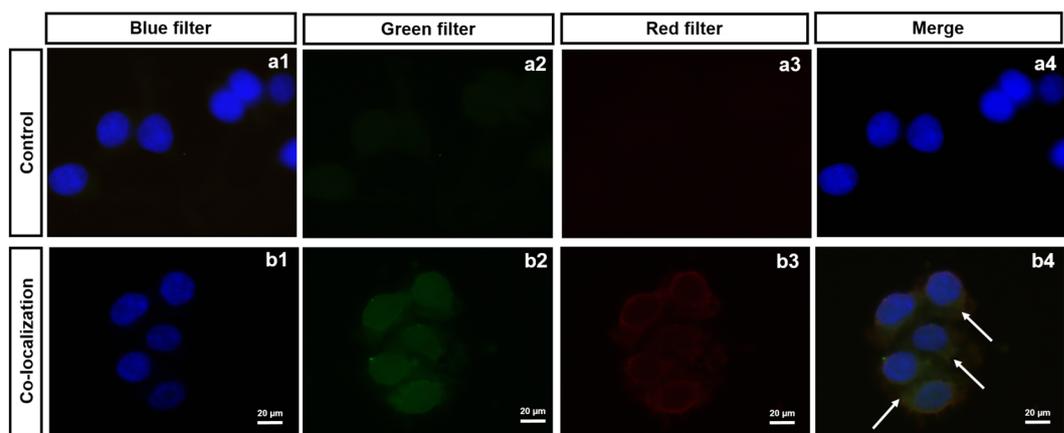
acids consisting in a disulfide constrained loop was used. This type of disulfide constrained peptide library have been proved useful for the selection of bioactive peptides, disease-specific antigen mimics and/or cell-specific peptides<sup>27–29</sup>.

During phage display, the phage pool between rounds was not amplified since it has been shown that the amplification decreases the library diversity, enriching clones that have an advantage during any of the amplification steps<sup>30,31</sup>. However, due to loss of phage titer, in the last two rounds the obtained phage particles were amplified using an engineered *E. coli* JM109<sup>+</sup> strain in order to minimize the presence of biased sequences<sup>19</sup>. The phage enrichment rate was progressively increased during the biopanning rounds, clearly indicating that phage clones with binding affinity to RKO cells were enriched (Table 1). This corroborates the expected process of affinity selection, where the specificity of the phage-libraries increases with additional rounds while the diversity drops<sup>30,32</sup>.

Several random clones from rounds 3, 4 and 5 were sequenced being the retrieved peptide sequences analyzed (Table 2). DNA sequencing showed that two peptide sequences appeared two times in sequencing results,



**Figure 7.** Effect of MCT1 and MCT4 downregulation in RKOpep binding. **(A)** Western blot analysis of MCT1 (47 kDa) and MCT4 (43 kDa) expression in siMCT1 and siMCT4 RKO cells. Cells were transfected with a negative control siRNA, siMCT1, siMCT4 and the expression of MCT isoforms was evaluated after 72 h. β-tubulin (50 kDa) was used to confirm an equal protein loading. Band intensity was measured using Image J software program. Similar results were obtained from three independent experiments and a representative blot is shown. **(B)** Analysis of binding ability of RKOpep and SCRpep to control, siMCT1 and siMCT4 RKO cells, assessed by flow cytometry.



**Figure 8.** Co-localization of RKOpep with MCT1 by immunocytochemistry. Control represents RKO cells incubated only with DAPI (a). Microscopy results expressing the co-localization of the highest concentration (50 μM) of FAM-RKOpep with an anti-MCT1 antibody conjugated with a secondary antibody coupled to Alexa Fluor Plus 680 to RKO cells (b). (1) blue, nuclei stained with DAPI, (2) green, FAM-RKOpep, (3) red, Alexa Fluor Plus 680 and (4) overlapping of all filters (1, 2 and 3). Scale bars represent 20 μm.

CPKSNNGVC and CKTPNGHLC. Afterwards, an *in vitro* cell-based ELISA assay was performed to assess specific binding of the phage clones from the last round of screening to RKO cells (Fig. 1). As a result, RKO-R5-1 phage clone encoding the sequence CPKSNNGVC was the best candidate with the highest binding ability being then selected for further studies.

Although biased sequences were minimized using *E. coli* JM109<sup>+</sup>, the peptide sequence CPKSNNGVC was analyzed to eliminate the possibility of the peptide being a TUP or a false positive (Supplementary Table S1). No relevant similarities were found using SAROTUP between the peptide here identified and those previously published, highlighting the novelty of this peptide.

As RKO-R5-1 phage clone, the free FAM-labelled peptide RKOpep also showed specific selectivity towards RKO cells in comparison to the CCD-841-CoN normal cells. Using microscopy, it was shown for the RKO cells that the fluorescence intensity increased with an increase of FAM-RKOpep concentration, which was not verified for the CCD-841-CoN cells (Fig. 2A). The cytometry analysis also corroborated these results, clearly showing that there is no affinity to the normal cells and that the RKOpep selectively bound to RKO cells (Fig. 2B). Moreover, the FAM-RKOpep showed noteworthy specificity for other three human CRC cells, HCT 116, Caco-2 and HCT-15 (Fig. 3). In addition, the peptide specific targeting of human colorectal cancer tissues, shown by immunofluorescence staining, revealed meaningful results (Fig. 4). Visible and noticeable green fluorescent signal of FAM-RKOpep was detected in human CRC tissues and a weak signal was observed for normal-adjacent ones. This experimental data supports the potential translation of this peptide to clinical oncology. The ability of peptides to recognize human cancer tissues was also demonstrated by Liu and collaborators<sup>33</sup>. They identified novel cancer-specific peptides selected through *in vivo* screening of phage display peptide libraries, that were able to target human cervical cancer tissues.

Identification of the cellular proteins responsible for peptide binding can lead to the discovery of central cellular targets previously unknown, not only by providing the information about the molecules expressed in the pathological state, but also by improving the understanding about what is not expressed under normal physiological conditions<sup>34</sup>. The results described above suggest that RKOpép recognizes a protein that is selectively overexpressed on the surface of at least four human CRC cell lines. Regarding this information, about 59 candidate proteins for the RKOpép were identified using a complete set of cell surface proteins of five CRC cell lines, including RKO, Caco-2 and HCT 116, described by de Wit and collaborators<sup>20</sup>.

To narrow this list of candidate proteins, a multiple sequence alignment for known CRC-related proteins was performed, leading to the identification of two potential targets (Supplementary Table S2). According to this homology search query, RKOpép was found to recognize MCT1, one of the 59 candidate proteins retrieved from the cell surface proteomic analysis after applying the selection criteria.

MCT1 catalyzes the proton-linked movement of many monocarboxylates, such as lactate and pyruvate, across the plasma membrane, and is the major transporter involved in lactate influx into tumor cells<sup>35,36</sup>. In addition to the important functional role of MCT1 in tumor metabolism, increased levels of MCT1 have been implicated in disease progression and prognosis in several human malignancies<sup>37–39</sup> including CRC<sup>40</sup>. MCT1 has a pivotal role in CRC maintenance and supports its use as a biomarker in primary and metastatic CRC<sup>41</sup>, which makes it a promising target for the novel RKOpép peptide herein identified.

MCT1 has a classical 12 transmembrane-helix structure with intracellular C- and N-terminus and a large cytosolic loop between transmembrane helices 6 and 7<sup>42</sup>. Although no 3D X-ray crystal structure is available for MCT1, its structure was predicted using PHYRE2. The docking results of RKOpép-MCT1 clearly indicated a predominance of hydrophobic interactions over hydrogen bonds (Fig. 5). The amino acid residues of RKOpép defined intensive hydrophobic interactions with the protein residues Met134, Pro74, Tyr70, Leu393, Leu128, Val397, Leu374, Ala150, Val373, Ile329, Met151, Asn147, Phe375, Ser154, Ala131 and Arg142. The majority of these hydrophobic interactions were established with amino acids located at the transmembrane helices of MCT1. In addition, the amino group of RKOpép Cys1 potentially formed a hydrogen bond with the MCT1 Asn127 carbonyl group at 2.70 Å. These preliminary bioinformatics analyses demonstrated that the RKOpép peptide can be specific to MCT1.

The expression level of MCT1 in normal and CRC cell lines was assessed by western blot (Fig. 6). The levels of expression of all CRC cell lines are in good agreement with the binding affinity of the RKOpép to them, as showed by the microscopy and cytometry results. Moreover, the specific recognition of human colorectal cancer tissues by the RKOpép can also be correlated with the overexpression of MCT1 at the plasma membrane of primary CRC tumors, as shown by the immunohistochemistry experiments performed by Martins and coworkers<sup>41</sup>.

To support our bioinformatics analysis, gene knockdown approaches (Fig. 7) and immunocytochemistry (Fig. 8) experiments were conducted using RKO cells. MCT1 silencing led to a significant decrease of the RKOpép binding to RKO cells as compared to the control, thus supporting that the MCT1 transmembrane protein is the target of RKOpép. The MCT family comprises a diverse group of transmembrane proteins. Two members of this family, MCT1 and MCT4, play key roles in the metabolic activity of tissues by mediating the proton-coupled transport of monocarboxylic acids across the plasma membrane<sup>43</sup>. Both MCT1 and MCT4 have been widely studied in colorectal cancer cells<sup>44,45</sup>.

To understand the binding selectivity of RKOpép to MCT1 membrane protein, the knockdown of MCT4 was also evaluated. A low reduction of RKOpép binding ability to RKO cells was observed, probably due to the fact that the siMCT1 had an impact in the MCT4 protein level as observed in the western blot (Fig. 7).

A co-localization of the FAM-labelled RKOpép and the anti-MCT1 antibody was seen, proven by the yellow/orange color obtained, thus confirming the affinity and specificity of the phage display selected peptide to the membrane protein MCT1 and therefore, attesting its usefulness for CRC diagnostics. Geng and collaborators also used immunocytochemistry staining to show that epidermal growth factor receptor 2 (HER2)-specific peptides, obtained through a combination of informatics approaches, including molecular dynamics modeling, co-localize with HER2<sup>46</sup>.

In conclusion, based on the binding of RKOpép to more than one CRC cell line and its ability to specifically recognize human colorectal cancer tissues, we reasoned that this peptide would be a valuable tool for CRC targeting, useful in new diagnostics and/or treatment approaches.

## Materials and Methods

**CRC primary tumor human samples.** Tissue samples from 67 patients treated in Hospital de Braga, North of Portugal, between 1<sup>st</sup> January of 2005 and 1<sup>st</sup> January of 2010 with CRC diagnosis were collected prospectively. Tumor localization was recorded and classified as colon and rectum (between anal verge and 15 cm at rigid rectoscopy). The histological type of CRC was classified by an experienced pathologist. CRC samples were included into tissue microarrays (TMAs). Prior to TMA construction, haematoxylin and eosin sections were reviewed to select representative areas of the tumor. Normal-adjacent tissue was also included in the TMAs for primary tumors. Each case was represented in the TMA by at least two cores of 0.6 mm. Tissue cores of kidney were used as controls for TMA orientation. The study protocol was approved by the Ethics Committee of Hospital de Braga. The data of CRC series was collected prospectively, patients were informed and signed a written consensus for collecting data and samples collection. All experiments were performed in accordance with relevant guidelines and regulations.

**Cell Lines and Cell Culture.** The human epithelial CRC cell line RKO (ATCC CRL 2577), normal colon cell line CCD-841-CoN (ATCC CRL 1790) and CRC cell line Caco-2 (ATCC HTB-37) were cultured in Dulbecco's Modified Eagle Medium (DMEM, Biochrom). The human CRC cell lines HCT 116 (ATCC CCL-247) and HCT-15 (ATCC CCL-225) were grown in Roswell Park Memorial Institute Medium 1640 (RPMI, Biochrom)

supplemented with 10% (v/v) fetal bovine serum (FBS, Biochrom) and 1% (v/v) penicillin-streptomycin (Biochrom). Cells were maintained at 37 °C with 5% CO<sub>2</sub>.

**BRASIL.** A commercial phage display library, Ph.D. – C7C, of seven random amino acids flanked by a pair of cysteine residues fused to the minor coat protein pIII of the M13KE phage from New England BioLabs (NEB), was used. An adaptation of the BRASIL method using RKO cells as target<sup>11</sup> was followed as described elsewhere<sup>27</sup>. The biopanning procedure was repeated four times to enrich the selected phages. A final negative selection step was performed with the normal CCD-841-CoN cell line.

**Selection and amplification of positive clones.** Single-stranded DNA (ssDNA) of individual phage clones from rounds 3, 4 and 5 was isolated according to a standard protocol described in<sup>27,47</sup> and amplified as detailed in<sup>48</sup>.

**DNA sequencing.** Before sequencing, the DNA was purified using the Illustra ExoProStar 1-Step kit (GE Healthcare). Sequencing was carried out by GATC Biotech using the M13-pIII primer 5'-TTAACTCCCTGCAAGCCTCA-3'. The SnapGene Version 1.1.3 (GSL Biotech) was used for peptide analysis. The corresponding peptide similarities were identified by Clustal Omega analysis<sup>49</sup>.

**Binding assays using ELISA.** RKO and CCD-841-CoN cells were plated in 96-well plates 24 h prior to the addition of positive phage clones at a concentration of  $1 \times 10^{10}$  PFU/well and incubated at 4 °C for 1 h. The plate was washed three times with PBS 1X with Tween 20 (PBS-T) prior to incubation with horseradish peroxidase (HRP)-conjugated anti-M13 monoclonal antibody (GE Healthcare), diluted in 1% of bovine serum albumin (BSA) in PBS 1X to the final dilution of 1:5000, and incubated at room temperature for 1 h. Subsequently, the plate was washed with PBS-T three times and the freshly prepared o-phenylenediamine dihydrochloride (OPD, Thermo Scientific) substrate for HRP detection was added to each well and incubated for 15 min. The plate was read at 450 nm on an automated ELISA plate reader (Biotech Synergy HT). The M13KE wild-type phage was used as a negative control. Triplicate measurements were performed at each data point and the signals obtained for each condition were compared.

**RKOpep.** The translated peptide sequence (RKOpep) from the phage DNA sequence giving the best binding results in ELISA was synthesized with a N-terminal modification (5/6 FAM), using the latest Fmoc solid-phase technology by Thermo Scientific Custom Peptide Synthesis Service. A scrambled peptide (SCRpep) was also synthesized as a negative control. Both peptides were HPLC purified to a purity >99%, lyophilized and analyzed using matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS) to confirm its molecular mass.

**Binding Assays.** For fluorescence experiments, RKO and CCD-841-CoN cells were seeded into plastic coverslips (SPL Life Sciences). After 24 h, cells were fixed with 4% paraformaldehyde (PFA, Sigma) in PBS 1X (w/v) at room temperature for 40 min. After rinsing with PBS 1X, cells were then blocked with PBS 1X containing 1% of BSA for 30 min. FAM-RKOpep was serially diluted in PBS 1X from 0 to 50 μM and incubated with human colorectal cells for 1 h at 4 °C. After being washed with PBS 1X, the coverslips were stained with Vectashield mounting media containing 4',6-diamidino-2-phenylindole dihydrochloride (DAPI, Vector Laboratories) solution at a concentration of 10 μg/mL, to counterstain the cell nucleic. The images were acquired by an Olympus BX51 microscope incorporated with a high-sensitivity camera Olympus DP71 at 10X magnification. For cytometry assays, a similar protocol was followed. RKO, CCD-841-CoN, Caco-2, HCT 116 and HCT-15 cells were scrapped and recovered for flow cytometry analysis performed with an Epics XL (Beckman Coulter) and at least 20 000 events were counted. Data were analyzed using the Flowing software (version 2.5.1).

**Immunofluorescence staining.** TMAs were used to confirm the ability of RKOpep to specifically recognize human CRC tissues. The slides were deparaffinized, rehydrated and after antigenic retrieval, that was performed by heating slides in 10 mM sodium citrate buffer pH 6.0 (Sigma) at 95 °C for 20 min and then slow cooling at room temperature for about 20 min, the tissue slides were blocked with TBS 1X with Tween 20 (TBS-T) for 30 min at 4 °C. Afterwards, the slides were incubated with FAM-RKOpep (50 μM) for 1 h at 4 °C and washed with TBS-T. Finally, the slides were stained with Vectashield mounting media containing DAPI solution. Images were acquired by an Olympus BX51 microscope incorporated with a high-sensitivity camera Olympus DP71 at 40X magnification.

**Bioinformatics analysis.** To determine the existence of potential false positives, target unrelated peptides (TUPs), multiple occurrences and already existing mimotopes, several web-based tools were used including BLAST and SAROTUP<sup>50</sup>. Next, to identify the receptor recognized by the peptide, high-resolution cell surface proteomics data was used<sup>20</sup>. The RKOpep peptide sequence was analyzed for homology to proteins with known or putative CRC correlations. The query was performed using the BLAST program (version BLAST+ 2.7.1) against the *Homo sapiens* non-redundant protein database using Blastp (BLASTp with word size of 3 and Blosum62 matrix, <http://blast.ncbi.nlm.nih.gov/>)<sup>51</sup>. Furthermore, molecular docking was performed to study peptide-protein interactions. The MCT1 structure used for peptide-docking analysis was predicted by homology modeling using PHYRE2 software<sup>24</sup>. Computational docking analysis, without a previous knowledge about the binding site, was performed using CABS-dock web server with the default parameters<sup>21</sup>. LigPlot+ program (version 1.4.5) was used to illustrate schematic two-dimensional representation of peptide-protein interaction<sup>52</sup>. PyMol (version 2.0.7) was employed for the display and analysis of molecular structures.

**Western blotting.** For western blot, about 20 µg of total protein were loaded, separated by 10% sodium dodecyl sulfate-polyacrylamide gel electrophoresis and transferred onto a nitrocellulose membrane (Protran Amersham, GE Healthcare). Then, to avoid non-specific interactions, the membrane was blocked with 5% BSA in TBS-T for 30 min at room temperature and then exposed overnight at 4 °C to the primary antibodies (dilution in TBS-T with 5% BSA) MCT1 (1:200 dilution, H-1, sc-365501, Santa Cruz Biotechnology), MCT4 (1:500 dilution, G-7, sc-376465, Santa Cruz Biotechnology), β-actin (1:2000 dilution, #4967S, Cell Signaling Technology) and β-tubulin (1:3000 dilution, 605101, Biolegend). After washing in TBS-T, the membranes were incubated with anti-mouse (sc-2031, Santa Cruz Biotechnology) or anti-rabbit (#7074S, Cell Signaling Technology) secondary antibodies linked to HRP (1:5000 dilution in TBS-T with 5% BSA) for 90 min at room temperature, respectively. The bands were detected using a chemiluminescent substrate (Clarity Western ECL Substrate, Bio-Rad) in ChemiDoc XRS + System (Bio-Rad). Band intensity was quantified by ImageJ.

**Downregulation of MCT1 and MCT4 expression.** Silencing of MCT1 and MCT4 expression was performed using 5 nM of Silencer Select Pre-Designed & Validated siRNAs from Thermo Fisher (MCT1 siRNA: s580 and MCT4 siRNA: s17417), using an adequate control (Silencer Select Negative Control siRNA: 4390843). Cells were transfected using Lipofectamine RNAiMAX Reagent (Thermo Fisher) according manufacturer's instructions.

**Immunocytochemistry for co-localization of FAM-RKOpep and MCT1.** For immunocytochemistry, approximately  $2 \times 10^5$  of RKO cells were seeded into plastic coverslips and cultured overnight. FAM-RKOpep was diluted in PBS 1X (50 µM) and incubated with adherent RKO cells for 1 h at 4 °C. After rinsing with PBS, the cells were fixed with 4% of PFA for 40 min and then permeabilized with 0.01% Triton-X (Merck) for 5 min. Afterwards, non-specific binding sites were blocked with TBS-T 1X containing 5% BSA for 30 min at 4 °C, and next exposed overnight at 4 °C to the primary antibody (dilution in TBS-T with 5% BSA) MCT1 (1:200 dilution, AB3538P, Merck). After rinsing with PBS 1X, the cells were incubated with anti-rabbit (#A-11011, Thermo Scientific) secondary antibody coupled to Alexa Fluor Plus 680 (1:500 dilution in TBS-T with 5% BSA) for 90 min at room temperature. After being washed with PBS 1X, the cells were stained with Vectashield mounting media containing DAPI. The images were acquired by an Olympus BX51 microscope incorporated with a high-sensitivity camera Olympus DP72 at 100X magnification.

**Statistical analysis.** Data was expressed as the mean ± standard deviation (SD) of three independent experiments. Two-way ANOVA with Tukey post-test was performed using GraphPad Prism (Graph-Pad Prism version 6.00 for Windows), to identify differences among multiple groups, considering a significance level of 95%.

## References

- Bray, F. *et al.* Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA. Cancer J. Clin.* **33**, 2597–2604 (2018).
- Grady, W. M. & Carethers, J. M. Genomic and Epigenetic Instability in Colorectal Cancer Pathogenesis. *Gastroenterology* **135**, 1079–1099 (2008).
- Gray, B. P. & Brown, K. C. Combinatorial peptide libraries: mining for cell-binding peptides. *Chem. Rev.* **114**, 1020–81 (2014).
- Gray, B. P., Li, S. & Brown, K. C. From Phage Display to Nanoparticle Delivery: Functionalizing Liposomes with Multivalent Peptides Improves Targeting to a Cancer Biomarker. *Bioconjug. Chem.* **24**, 85–96 (2013).
- Paschke, M. Phage display systems and their applications. *Appl. Microbiol. Biotechnol.* **70**, 2–11 (2006).
- Sidhu, S. S. Engineering M13 for phage display. *Biomol. Eng.* **18**, 57–63 (2001).
- Budynek, P., Dąbrowska, K., Skaradziński, G. & Górski, A. Bacteriophages and cancer. *Arch. Microbiol.* **192**, 315–320 (2010).
- Ferreira, D. & Martins, I. M. Artificial virus particles. In *Bioinspired Materials for Medical Applications* 427–450, <https://doi.org/10.1016/B978-0-08-100741-9.00015-2> (Elsevier, 2017).
- Bar, H., Yacoby, I. & Benhar, I. Killing cancer cells by targeted drug-carrying phage nanomedicines. *BMC Biotechnol.* **8**, 37 (2008).
- Smith, G. P. Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science* **228**, 1315–7 (1985).
- Giordano, R. J., Cardó-Vila, M., Lahdenranta, J., Pasqualini, R. & Arap, W. Biopanning and rapid analysis of selective interactive ligands. *Nat. Med.* **7**, 1249–53 (2001).
- de Oliveira, H. C. *et al.* Peptides Derived from a Phage Display Library Inhibit Adhesion and Protect the Host against Infection by *Paracoccidioides brasiliensis* and *Paracoccidioides lutzii*. *Front. Pharmacol.* **7**, 509 (2016).
- Jin, W. *et al.* Discovery of PSMA-specific peptide ligands for targeted drug delivery. *Int. J. Pharm.* **513**, 138–147 (2016).
- Dobroff, A. S. *et al.* Towards a transcriptome-based theranostic platform for unfavorable breast cancer phenotypes. *Proc. Natl. Acad. Sci.* **113**, 12780–12785 (2016).
- Reetz, J. *et al.* Novel subventricular zone early progenitor cell-specific adenovirus for *in vivo* therapy of central nervous system disorders reinforces brain stem cell heterogeneity. *Brain Struct. Funct.* **221**, 2049–2059 (2016).
- Krag, D. N. *et al.* Selection of Tumor-binding Ligands in Cancer Patients with Phage Display Libraries. *Cancer Res.* **66**, 7724–7733 (2006).
- Petrenko, V. A. & Jayanna, P. K. Phage protein-targeted cancer nanomedicines. *FEBS Lett.* **588**, 341–349 (2014).
- Huang, J., Ru, B. & Dai, P. Bioinformatics resources and tools for phage display. *Molecules* **16**, 694–709 (2011).
- Umlauf, B. J., McGuire, M. J. & Brown, K. C. BioTechniques - Introduction of plasmid encoding for rare tRNAs reduces amplification bias in phage display biopanning. *Biotechniques* **58**, 81–4 (2015).
- de Wit, M. *et al.* Cell surface proteomics identifies glucose transporter type 1 and prion protein as candidate biomarkers for colorectal adenoma-to-carcinoma progression. *Gut* **61**, 855–864 (2012).
- Kurcinski, M., Jamroz, M., Blaszczyk, M., Kolinski, A. & Kmiecik, S. CABS-dock web server for the flexible docking of peptides to proteins without prior knowledge of the binding site. *Nucleic Acids Res.* **43**, W419–W424 (2015).
- Blaszczyk, M. *et al.* Modeling of protein-peptide interactions using the CABS-dock web server for binding site search and flexible docking. *Methods* **93**, 72–83 (2016).
- Ciemny, M. P., Kurcinski, M., Kozak, K. J., Kolinski, A. & Kmiecik, S. Highly Flexible Protein-Peptide Docking Using CABS-Dock. In *Methods in molecular biology (Clifton, N.J.)* **1561**, 69–94 (2017).
- Kelley, L. A. & Sternberg, M. J. E. Protein structure prediction on the Web: a case study using the Phyre server. *Nat. Protoc.* **4**, 363–371 (2009).

25. American Cancer Society. Cancer Facts & Figures 2014. *Atlanta Am. Cancer Soc.* (2014).
26. Mouradov, D. *et al.* Colorectal Cancer Cell Lines Are Representative Models of the Main Molecular Subtypes of Primary Cancer. *Cancer Res.* **74**, 3238–3247 (2014).
27. Silva, V. L. *et al.* Selection of Novel Peptides Homing the 4T1 CELL Line: Exploring Alternative Targets for Triple Negative Breast Cancer. *PLoS One* **11**, e0161290 (2016).
28. Sakamoto, K. *et al.* Identification of ligand-selective peptidic ActRIIB-antagonists using phage display technology. *Biochem. Biophys. Reports* **11**, 33–39 (2017).
29. Mintz, P. J. *et al.* Discovery and horizontal follow-up of an autoantibody signature in human prostate cancer. *Proc. Natl. Acad. Sci.* **112**, 2515–2520 (2015).
30. Derda, R. *et al.* Diversity of phage-displayed libraries of peptides during panning and amplification. *Molecules* **16**, 1776–803 (2011).
31. Vodnik, M., Zager, U., Strukelj, B. & Lunder, M. Phage Display: Selecting Straws Instead of a Needle from a Haystack. *Molecules* **16**, 790–817 (2011).
32. Christiansen, A. *et al.* High-throughput sequencing enhanced phage display enables the identification of patient-specific epitope motifs in serum. *Sci. Rep.* **5**, 12913 (2015).
33. Liu, X. *et al.* Selection and identification of novel peptides specifically targeting human cervical cancer. *Amino Acids* **50**, 577–592 (2018).
34. Wu, C.-H., Liu, I.-J., Lu, R.-M. & Wu, H.-C. Advancement and applications of peptide phage display technology in biomedical science. *J. Biomed. Sci.* **23**, 8 (2016).
35. Sprowl-Tanio, S. *et al.* Lactate/pyruvate transporter MCT-1 is a direct Wnt target that confers sensitivity to 3-bromopyruvate in colon cancer. *Cancer Metab.* **4**, 20 (2016).
36. Bola, B. M. *et al.* Inhibition of Monocarboxylate Transporter-1 (MCT1) by AZD3965 Enhances Radiosensitivity by Reducing Lactate Transport. *Mol. Cancer Ther.* **13**, 2805–2816 (2014).
37. Miranda-Gonçalves, V. *et al.* Monocarboxylate transporters (MCTs) in gliomas: expression and exploitation as therapeutic targets. *Neuro. Oncol.* **15**, 172–188 (2013).
38. Pinheiro, C. *et al.* Monocarboxylate transporter 1 is up-regulated in basal-like breast carcinoma. *Histopathology* **56**, 860–867 (2010).
39. Pinheiro, C. *et al.* Increasing Expression of Monocarboxylate Transporters 1 and 4 Along Progression to Invasive Cervical Carcinoma. *Int. J. Gynecol. Pathol.* **27**, 568–574 (2008).
40. Pinheiro, C. *et al.* Increased expression of monocarboxylate transporters 1, 2, and 4 in colorectal carcinomas. *Virchows Arch.* **452**, 139–146 (2008).
41. Martins, S. F. *et al.* Significance of glycolytic metabolism-related protein expression in colorectal cancer, lymph node and hepatic metastasis. *BMC Cancer* **16**, 535 (2016).
42. Halestrap, A. P. The monocarboxylate transporter family-Structure and functional characterization. *IUBMB Life* **64**, 1–9 (2012).
43. Park, S. J. *et al.* An overview of MCT1 and MCT4 in GBM: small molecule transporters with large implications. *Am. J. Cancer Res.* **8**, 1967–1976 (2018).
44. Amorim, R. *et al.* Monocarboxylate transport inhibition potentiates the cytotoxic effect of 5-fluorouracil in colorectal cancer cells. *Cancer Lett.* **365**, 68–78 (2015).
45. Ferro, S. *et al.* Characterization of acetate transport in colorectal cancer cells and potential therapeutic implications. *Oncotarget* **7** (2016).
46. Geng, L. *et al.* Structure-based Design of Peptides with High Affinity and Specificity to HER2 Positive Tumors. *Theranostics* **5**, 1154–65 (2015).
47. NEB. Ph.D.TM Phage Display Libraries Manual.
48. Nobrega, F. L. *et al.* Screening and characterization of novel specific peptides targeting MDA-MB-231 claudin-low breast carcinoma by computer-aided phage display methodologies. *BMC Cancer* **16**, 881 (2016).
49. Sievers, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539–539 (2014).
50. Huang, J., Ru, B., Li, S., Lin, H. & Guo, F.-B. SAROTUP: Scanner and Reporter of Target-Unrelated Peptides. *J. Biomed. Biotechnol.* **2010**, 1–7 (2010).
51. Altschul, S. F. *et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–402 (1997).
52. Laskowski, R. A. & Swindells, M. B. LigPlot+: Multiple Ligand-Protein Interaction Diagrams for Drug Discovery. *J. Chem. Inf. Model.* **51**, 2778–2786 (2011).

## Acknowledgements

The authors acknowledge Dr. Mariana Henriques from the CEB for providing the Caco-2 cell line, Dr. Manuela Côrte-Real from the Center for Molecular and Environmental Biology (CBMA) for supplying the HCT 116 and HCT-15 cells and Diana Vilas-Boas from CEB for technical assistance in microscopy assays. This study was supported by the Portuguese Foundation for Science and Technology (FCT) under the scope of the strategic funding of UID/BIO/04469/2019 unit and BioTecNorte operation (NORTE-01-0145-FEDER-000004) funded by the European Regional Development Fund under the scope of Norte2020 - Programa Operacional Regional do Norte and the Project FCOMP-01-0124-FEDER-021053 (PTDC/SAU-BMA/121028/2010). Débora Ferreira is recipient of a fellowship supported by a doctoral advanced training (call NORTE-69-2015-15) funded by the European Social Fund under the scope of Norte2020 - Programa Operacional Regional do Norte. Franklin L. Nobrega, Sara Granja and Ligia R Rodrigues acknowledge FCT for the grants SFRH/BD/86462/2012, SFRH/BPD/117858/2016 and SFRH/BSAB/142991/2018, respectively. Catarina Barbosa-Matos also acknowledge her research grant UMINHO/BI/395/2018.

## Author Contributions

L.R. conceived the study. D.F. and A.S. performed the experiments. D.F., A.S. and I.M. analyzed the data and wrote de manuscript. L.R., F.B., I.M., C.B.M., S.G. and F.N. gave suggestions on the experimental design. S.F.M. performed C.R.C. surgery. All authors reviewed the results and approved the final version of the manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-019-40562-1>.

**Competing Interests:** The authors declare that the research was conducted in the absence of any commercial, financial and non-financial relationships that could be construed as a potential conflict of interest.

**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