CANCER CELL TARGETING PEPTIDE

TECHNICAL FIELD

The present disclosure relates to homing peptides. More specifically, the disclosure relates to homing peptides that can target cancerous cells, and their useful applications in therapy and imaging, including peptides with the ability to cross the healthy blood-brain-barrier (BBB) and methods of use thereof for targeting cancerous cells from the circulation into the brain and other organs/tissues containing cancerous cells.

BACKGROUND

This section illustrates useful background information without admitting any technique described herein as being representative of the state of the art.

Diagnostic imaging and therapeutic pharmaceuticals play an important role in modern medicine. Some applications have been developed that involve use of a homing peptide which specifically binds to a selected target, such as a cancerous cell or a metastatic cell.

Brain metastasis (BM) are late-stage complications of cancer commonly originating from melanoma, breast carcinomas, and lung carcinomas, but any cancer can metastasize to brain. Following an improved cancer patient survival due to better systemic therapies, the incidence of brain metastasis is increasing. Moreover, therapy of BM faces additional challenges such as i) the lack of specificity and selectivity to target tumour cells, ii) the inability of the drugs to cross the BBB, and iii) the chemoresistance/radioresistance of the tumour. BMs are associated with the terminal stage of cancer and due to the technical challenge of removing highly invasive tumour masses in the brain, BM patients are most of the time not eligible for neurosurgery or other types of therapy. In addition, metastatic lesions in other organs may be very difficult to diagnose or treat affecting the quality or length of life of the patients.

Peptide-drug conjugates are personalized medicine tools developed to optimize the diagnostic or treatment delivery to the tumour or any other selected target. Addition of a cancer-targeting moiety to the chemotherapy increases the drug specificity and delivery to the tumour site. Thus, decreasing the possible siden effect in healthy tissues. In addition, targeting peptides are used in the clinics as molecular probes for detection and clinical diagnosis of cancer.

Therefore, there is a need for novel targeting moieties that can be used to detect cancerous cells and tissues, and that can be used to deliver imaging agents for diagnostic purposes and therapeutic agents to cancerous/metastatic cells.

SUMMARY

The scope of protection sought for various embodiments of the invention is set out by the independent claims. Any example or embodiment which is described herein but which does not fall within the scope of the independent claims is presented herein as an additional example or additional embodiment which is useful for understanding the claimed invention.

The inventors identified a novel homing peptide CGRTLGRVA (SEQ ID NO: 4, here after referred to as RVA) and its functional variants by using an *ex vivo/in vivo* phage display screening. The novel homing peptide and its functional variants were found to specifically target the brain metastases, which property can be used to an advantage in various technical applications, such as in imaging, therapy, delivery of agents to cells, as a research tool, and in development of further therapeutic or diagnostic agents. In the present disclosure the term homing peptide is used when referring to RVA, and unless otherwise explicitly mentioned, the term encompasses all functional variants of the homing peptide. A functional variant is a chemically modified version of the homing peptide that has been specifically modified to have a different ability to target and home in on a specific cell or tissue, preferably to the same cell or tissue as the homing peptide. The chemical modification may be e.g. a substitution, addition, or deletion of an amino acid of the homing peptide, cyclization of the homing peptide, a chemical modification of the homing peptide, or an addition of a peptide, protein, domain, or another chemical moiety to the homing peptide. Properties of the functional variant, such as ability to home specifically to a target cell or tissue and an ability to pass BBB, can be measured as for the homing peptide, and compared with the properties of the homing peptide to verify functionality.

The potential of RVA as a delivery vehicle of targeted imaging agents and drugs for BM diagnostics and treatment was evaluated experimentally. The experiments show that an imaging agent comprising RVA and fluorine 18 was able to detect brain metastatic cells with high precision in PET imaging of preclinical animal models.

As a proof-of-concept for drug delivery, an antimicrobial peptide KLAKLAK was conjugated to the N-terminus of RVA. The RVA-drug conjugates killed all tested cancer cell lines with IC_{50} values in low micromolar range. Negligible or no cytotoxicity was observed in the non-cancer cell lines when using this conjugate.

Moreover, peroxiredoxin 1 (PRDX1) was identified as the RVA target protein by using photoaffinity labelling and mass spectrometry (MS). PRDX1-RVA binding was validated by molecular interaction studies using the Surface Plasmon Resonance (SPR) and Microscale Thermophoresis (MST) assays. In addition, alanine scanning technique was used to determine the effect of each amino acid in the RVA sequence on binding to the recombinant human PRDX1. High PRDX1 expression was detected in preclinical brain metastases models, in clinical BM samples, metastatic lesions in bones and lungs as well as in multiple primary tumours (Examples 9 and 10).

According to a first example aspect is provided a homing peptide, or a functional variant thereof, comprising, or consisting of, an amino acid sequence CxxxLG, SEQUENCE ID NO: 1.

In an embodiment the present homing peptide or a functional variant thereof preferably comprises, or consists of, an amino acid sequence CxxxLGxxx, SEQUENCE ID NO: 2.

In an embodiment the present homing peptide or a functional variant thereof preferably comprises, or consists, of an amino acid sequence CxRxLGRVA, SEQUENCE ID NO: 3.

In an embodiment the present homing peptide or a functional variant thereof preferably comprises, or consists of, an amino acid sequence CGRTLGRVA, SEQ ID NO: 4.

In an embodiment the present homing peptide or a functional variant thereof preferably comprises, or consists of, an amino acid sequence CxRxLG, SEQ ID NO: 5.

In an embodiment the present homing peptide or a functional variant thereof preferably comprises, or consists of, an amino acid sequence CGRTLG, SEQ ID NO: 6.

In an embodiment the present homing peptide or a functional variant thereof preferably comprises, or consists of, an amino acid sequence CGxxLGxxA, SEQ ID NO: 46 or CxxxLGxxA, SEQ ID NO: 47 or CxxxLGRVA, SEQ ID NO: 48.

In an embodiment the present homing peptide or a functional variant thereof preferably comprises, or consists of, an amino acid sequence CGxxLGRVA, SEQ ID NO: 49 or CxxxLGRVA, SEQ ID NO: 50.

In an embodiment the present homing peptide or a functional variant thereof preferably comprises, or consists of, an amino acid sequence CGxxLG, SEQ ID NO: 51 or CGxxLGxxx, SEQ ID NO: 52 or CGxxLGRVA, SEQ ID NO: 49.

In an embodiment the present homing peptide or a functional variant thereof preferably comprises, or consists of, an amino acid sequence CGRxLG, SEQ ID NO: 53, or CGRxLGxxx, SEQ ID NO: 54, or CGRxLGRVA, SEQ ID NO: 55.

In an embodiment the present homing peptide or a functional variant thereof has no more than 30 amino acids. Preferably the sequence of SEQ ID NO: 1-6 is in the N-terminus or in the C-terminus of the homing peptide or its functional variant, or it is flagged by N-terminal and C-terminal sequences.

In an embodiment the present homing peptide or a functional variant thereof has a molecular weight selected from a range 1000-1500g/mol, 1000-1450 g/mol, 1100-1500 g/mol, 1100-1450 g/mol, 1200-1500g/mol, and 1200-1450 g/mol. In an embodiment the present homing peptide or a functional variant thereof is capable of specifically binging PRDX1 (peroxiredoxin 1), preferably the N-terminal domain of PRDX1, and/or optionally to a corresponding protein in another organism. The term PRDX1 refers to the established gene name of the human peroxiredoxin 1, and the N-terminal domain to its N-terminal domain as identified in the PRDX1 structure. The skilled person is able to identify PRDX1 and its N-terminal domain in relevant gene and protein sequence databases, such as via UniProt entry Q06830, www.uniprot.org.

In an embodiment the present homing peptide or a functional variant thereof is capable of crossing a healthy blood-brain-barrier (BBB).

In an embodiment the present homing peptide or a functional variant thereof is capable of being internalized by a cell expressing PRDX1.

According to an example aspect is provided a direct or indirect conjugate comprising the present homing peptide and a cargo molecule.

In the present disclosure, unless otherwise specifically mentioned, the term conjugate encompasses direct conjugates and indirect conjugates.

In an embodiment the cargo molecule is at least one of: a nucleic acid, antisense molecule, aptamer, ribozyme, triplex forming oligonucleotide, external guide sequence, RNAi, CRISPR/Cas, zinc finger nuclease, transcription activator-like effector nucleases, an expression vector encoding a protein or a functional nucleic acid, a vectors suitable for integration into a cells genome or expressed extra-chromosomally, antibody, monoclonal antibody, tagged molecule, fluorescence tag molecule, label, nanoparticle, conjugated nanoparticle, liposomal payload, small molecule, anti-cancer agent, an anti-inflammatory agent, an immunomodulatory agent, an antimicrobial agent, an organic compound, an organometallic compound; a hydrophilic compound, a hydrophobic compound, an amphiphilic compounds, toxic agent, radioisotope, a radionuclide, a fluorescent tag/compound/dye, a magnetic tag/molecule, paramagnetic molecule, x-ray imaging agent,

contrast media, dye, near infra-red dyes, SPECT imaging agent, PET imaging agent, peptide, lipid, glycolipid, glycoprotein, polypeptide, a therapeutic agent, a diagnostic agent, a prophylactic agent, a nutraceutical agent.

In an embodiment the nucleic acid is an antisense molecule, aptamer, ribozyme, triplex forming oligonucleotide, external guide sequence, ribonucleotide, RNAi, CRISPR/Cas.

In an embodiment the cargo molecule is at least one of: a hydrophilic agent, hydrophobic agent, amphiphilic compound, toxic agent, radioisotope, dye, a radionuclide, a fluorescent tag/compound/dye, a magnetic tag/molecule, paramagnetic molecule, x-ray imaging agent, contrast media, near infra-red dye, SPECT imaging agent, PET imaging agent, peptide, lipid, glycolipid, glycoprotein, polypeptide.

In an embodiment the cargo molecule is at least one of: a therapeutic agent, a diagnostic agent, a prophylactic agent, a nutraceutical agent.

In an embodiment the cargo molecule includes one or more therapeutic agents, such as two, three, or four therapeutic agents.

In an embodiment the nucleic acid is an expression vector encoding a protein or a functional nucleic acid.

In an embodiment the vector is suitable for integration into a cells genome or for being expressed extra-chromosomally by the cell.

In an embodiment the small molecule is an anti-cancer agent, an anti-inflammatory agent, an immunomodulatory agent, or an antimicrobial agent.

In an embodiment the cargo molecule is a therapeutic moiety.

In an embodiment the cargo molecule is an imaging moiety, or a diagnostic moiety, alone or in combination with one or more therapeutic moiety.

In an embodiment the cargo molecule is an imaging agent.

According to an example aspect is provided a pharmaceutical composition comprising the present conjugate and a pharmaceutically acceptable carrier.

In an embodiment in the cargo molecule of the conjugate, the cargo molecule is encapsulated within or conjugated to a linker or carrier. In an embodiment the linker or the carrier encapsulates or is complexed with the one or more active agent or cargo molecule. In an embodiment the linker or the carrier is directly or indirectly conjugated to the homing peptide. One or more cargo molecules can be conjugated or complexed with the homing

peptide by one or more linkers. In an embodiment the linker is cleavable. In another embodiment the linker is non-cleavable.

According to an example aspect is provided the present homing peptide, or the present conjugate, for use as a medicine. In an embodiment the conjugate is a direct conjugate. In another embodiment the conjugate is an indirect conjugate. In an embodiment a final formula comprising the homing peptide or the conjugate is suitable for mucosal, pulmonary, intranasal, oral, intravenous, intratumoural, or intramuscular delivery.

According to an example aspect is provided the present homing peptide, or the present conjugate, for treating or for in use of treating cancer, preferably a brain metastasis. In an embodiment the conjugate is a direct conjugate. In another embodiment the conjugate is an indirect conjugate. In an embodiment the use comprises treating a subject with an effective amount of the homing peptide and/or the conjugate.

According to an example aspect is provided a method of *in vivo* or *ex vivo* or *in vitro* diagnosis of cancer using the present homing peptide, or the present conjugate, wherein the cancer is preferably at least one of melanoma, breast carcinoma, lung carcinoma, brain cancer, brain metastatic cell, or any other cancer with PRDX1 expression. In an embodiment the conjugate is a direct conjugate. In another embodiment the conjugate is an indirect conjugate.

In an embodiment the cancer or metastatic cell is at least one of: AML, Bladder, Breast, Colon, Esophagus, Liver, Lung adenocarcinoma, Lung squamous carcinoma, Ovary, Pancreas, Prostate, Rectum, Skin, Stomach, testis, bladder, Uterus (CS and EC), or as presented in Table 7.

According to an example aspect is provided a method of delivering a cargo molecule into a cancer cell, comprising contacting the cancer cell with the present conjugate. In an embodiment the conjugate is a direct conjugate. In another embodiment the conjugate is an indirect conjugate. In an embodiment the method is an *in vivo* method. In another embodiment the method is an *in vitro* method.

According to an example aspect is provided a method of identifying a cancerous target cell or cancerous target tissue in a biological sample, the method comprising:

(a) contacting the biological sample with the present conjugate for a time and under conditions sufficient for specifically binding the conjugate to the target cell or the target tissue; and

(b) detecting the presence or absence of the bound conjugate in the biological sample; and

wherein presence of the conjugate in detecting step (b) indicates that the cancerous target cell or the cancerous target tissue is present in the biological sample.

According to an example aspect is provided a method of identifying a cancerous target cell or cancerous target tissue in a subject, the method comprising:

(a) contacting the subject, or a tissue of the subject, or a cell of the subject, with the present conjugate for a time and under conditions sufficient for specifically binding the conjugate to the target cell or the target tissue; and

(b) detecting the presence or absence of the bound conjugate in the subject; and

wherein presence of the conjugate in detecting step (b) indicates that the cancerous target cell or the cancerous target tissue is present in the subject.

In an embodiment the conjugate of the method is a direct conjugate. In another embodiment the conjugate is an indirect conjugate.

In an embodiment the step (b) is carried out by immunohistochemistry, fluorescent imaging, radioimmunoassay, and/or immunofluorescence.

According to an example aspect is provided a kit comprising: (a) a container comprising a pharmaceutical composition containing the present homing peptide, the present conjugate, or the present pharmaceutical composition, in solution or in a lyophilized form; (b) optionally, a second container containing a diluent or reconstituting solution for the lyophilized formulation; and (c) optionally, instructions for use of the solution or reconstitution and/or use of the lyophilized formulation.

According to another aspect is provided a complex comprising the present homing peptide and a target molecule. In an embodiment the target molecule is a molecule to which the homing peptide binds specifically, such as PRDX1, or an N-terminal domain of PRDX1.

Different non-binding example aspects and embodiments have been illustrated in the foregoing. The embodiments in the foregoing are used merely to explain selected aspects or steps that may be utilized in different implementations. Some embodiments may be presented only with reference to certain example aspects. It should be appreciated that corresponding embodiments may apply to other example aspects, as well. Moreover, in the following further embodiments are described that can be combined with the aspects and embodiments described above.

BRIEF DESCRIPTION OF THE FIGURES

Some example embodiments will be described with reference to the accompanying figures, in which:

Figures 1A-1L show results from experiments described in Example 1.

Figures 2A-2F show results from experiments described in Example 2.

Figures 3A-3L show results from experiments described in Example 3.

Figures 4A-4L show results from experiments described in Example 4.

Figures 5A-5M show results from experiments described in Example 7.

Figures 6A-6M show results from experiments described in Example 8.

Figure 7 shows results from experiments described in Example 9.

Figure 8 shows results from experiments described in Example 10.

Figures 9A-9B show results from experiments described in Example 2.

Figures 10A-10B show results from experiments described in Example 4.

Figures 10a-10d show results from experiments described in Example 4.

Figures 11A-11R shows results from experiments described in Example 5.

Figure12 shows results from experiments described in Example 8.

Figures 13A, 13B show results from experiments described in Example 11.

Figures 14A, 14B, 14C, 14D, 14E, and 14F show results from experiments described in Example 12.

Figure 15 shows results from experiments described in Example 13.

Figure 16 shows results from experiments described in Example 14.

Figures 17A, 17B, 17C, and 17D show results from experiments described in Example 15.

Figure 18 shows results from experiments described in Example 16.

Figure 19 shows results from experiments described in Example 17.

Figure 20 shows results from experiments described in Example 18.

SEQUENCE LISTINGS

SEQ ID NO: 1 is amino acid sequence CxxxLG.

SEQ ID NO: 2 is amino acid sequence CxxxLGxxx.

SEQ ID NO: 3 is amino acid sequence CxRxLGRVA.

SEQ ID NO: 4 is amino acid sequence CGRTLGRVA.

SEQ ID NO: 5 is amino acid sequence CxRxLG.

SEQ ID NO: 6 is amino acid sequence CGRTLG.

SEQ ID NO: 7 is amino acid sequence KLAKLAK.

SEQ ID NO: 8 is amino acid sequence CVAALNADG, control peptide.

SEQ ID NO: 9 is amino acid sequence AGRTAARVA, control peptide TrA Ctrl.

SEQ ID NO: 10 is amino acid sequence CAGTECLTC, LTC.

SEQ ID NO: 11 is amino acid sequence CGRTLGRVA, RVA.

SEQ ID NO: 12 is amino acid sequence CGRTRG, TRG.

SEQ ID NO: 13 is amino acid sequence CSHVSN, VSN.

SEQ ID NO: 14 is amino acid sequence CKLLSART, ART.

SEQ ID NO: 15 is amino acid sequence AGRTLGRVA, [Ala1] RVA.

SEQ ID NO: 16 is amino acid sequence CARTLGRVA, [Ala2] RVA.

SEQ ID NO: 17 is amino acid sequence CGATLGRVA, [Ala3] RVA.

SEQ ID NO: 18 is amino acid sequence CGRALGRVA, [Ala4] RVA.

SEQ ID NO: 19 is amino acid sequence CGRTAGRVA, [Ala5] RVA.

SEQ ID NO: 20 is amino acid sequence CGRTLARVA, [Ala6] RVA.

SEQ ID NO: 21 is amino acid sequence CGRTLGAVA, [Ala7] RVA.

SEQ ID NO: 22 is amino acid sequence CGRTLGRAA, [Ala8] RVA.

SEQ ID NO: 23 is amino acid sequence AVRGLTRGC, retroinverse-RVA.

SEQ ID NO: 24 is amino acid sequence CGRTLGRVA, [D-Cys1] RVA.

SEQ ID NO: 25 is amino acid sequence CGRTLGRVA, [D-Arg3] RVA.

SEQ ID NO: 26 is amino acid sequence CGRTLGRVA, [D-Thr4] RVA.

SEQ ID NO: 27 is amino acid sequence CGRTLGRVA, [D-Leu5] RVA.

SEQ ID NO: 28 is amino acid sequence CGRTLGRVA, [D-Arg7] RVA.

SEQ ID NO: 29 is amino acid sequence CGRTLGRVA, [D-Val8] RVA.

SEQ ID NO: 30 is amino acid sequence CGRTLGRVA, [D-Ala9] RVA.

SEQ ID NO: 31 is amino acid sequence CGRTLGRVC, Cyclic 'native' RVA.

SEQ ID NO: 32 is amino acid sequence CGRTLGRVAK, RVAK.

SEQ ID NO: 33 is amino acid sequence CVAALNADG, Negative control.

SEQ ID NO: 34 is amino acid sequence TAMRAACGRTLGRVA, TAMRA-A-RVA.

SEQ ID NO: 35 is amino acid sequence TAMRACGRTLGRVA, TAMRA-RVA.

SEQ ID NO: 36 is amino acid sequence TAMRACVAALNADG, TAMRA-Negative control.

SEQ ID NO: 37 is amino acid sequence ACGRTLGRVA, Biotin-Bpa-A-RVA.

SEQ ID NO: 38 is amino acid sequence ACGRTLGRVA, Biotin-ZpN3-A-RVA.

SEQ ID NO: 39 is amino acid sequence CVAALNADG, Biotin-Negative control.

SEQ ID NO: 40 is amino acid sequence AGRTAARVA, TrA Negative control.

SEQ ID NO: 41 is amino acid sequence KLAKLAKKLAKLAKGGCGRTLGRVA, KLAK-RVA.

SEQ ID NO: 42 is amino acid sequence D(KLAKLAKKLAKLAK), all D amino acids, Free KLAK.

SEQ ID NO: 43 is amino acid sequence MMAECGRTLGRVA, MMAE-RVA.

SEQ ID NO: 44 is amino acid sequence MMAEAGRTAARVA, KLAK-TrA Negative control.

SEQ ID NO: 45 is amino acid sequence MMAE, Free MMAE.

SEQ ID NO: 46 is amino acid sequence CGxxLGxxA.

SEQ ID NO: 47 is amino acid sequence CxxxLGxxA.

SEQ ID NO: 48 is amino acid sequence CxxxLGRVA.

SEQ ID NO: 49 is amino acid sequence CGxxLGRVA.

SEQ ID NO: 50 is amino acid sequence CxxxLGRVA.

SEQ ID NO: 51 is amino acid sequence CGxxLG.

SEQ ID NO: 52 is amino acid sequence CGxxLGxxx.

SEQ ID NO: 53 is amino acid sequence CGRxLG.

SEQ ID NO: 54 is amino acid sequence CGRxLGxxx.

SEQ ID NO: 55 is amino acid sequence CGRxLGRVA.

SEQ ID NO: 56 is amino acid sequence DACSRFLGERVDATAAGCSR.

SEQ ID NO: 57 is amino acid sequence CARRLGRVATTYYMDVW.

SEQ ID NO: 58 is amino acid sequence CARRLGRVARPTTWTS.

SEQ ID NO: 59 is amino acid sequence CGVRLGC.

SEQ ID NO: 60 is amino acid sequence CVPELGHEC.

The character “X” in amino acid sequences denotes any amino acid.

DETAILED DESCRIPTION

The terms “a” and “an” and “the” and similar in the context of describing features, elements, examples, or claims are to be construed to cover both the singular and the plural, unless otherwise indicated or clearly contradicted by context.

Recitation of ranges of values herein are merely intended to serve as a way of referring individually to each separate value falling within the range, unless otherwise indicated herein. Each separate value is thus disclosed in the specification as if it were individually recited. In an embodiment, and as is understood by the skilled person in the context in which the expression is used, an open-ended range such as “less than 10” is to be construed as a range disclosing values below 10, but above 0. Such a range can be understood to include a lower limit of e.g. 0.0001, 0.001, 0.1, or 1. If a lower limit is not recited, a lower limit of an open-ended range can be determined by the skilled person such that at least one technical effect is observable or measurable, thereby excluding insignificant trace amounts in the relevant context.

All methods described herein can be performed in any suitable order unless otherwise indicated herein or otherwise clearly contradicted by context. The use of any example, or exemplary language (e.g., “such as”, “for example”, and “optionally”) provided herein, is intended merely to better illustrate various embodiments and it does not limit the scope of the claims unless otherwise stated. No language in the specification should be construed as indicating that any non-claimed element is an essential feature of the invention.

In another embodiment the method steps are carried out in the sequence identified in any aspect, embodiment, example, or claim. In another embodiment any method step specified to be carried out to a product or an intermediate obtained in a preceding process step is carried out directly to said product or intermediate, i.e. without additional, optional, or auxiliary processing steps that may chemically and/or physically alter the product or intermediate between said two consecutive steps. For example, a washing step or extraction of material typically alters the chemical composition of the material in a chemical process.

Unless otherwise indicated, all numbers expressing quantities of ingredients, reaction conditions, and so forth used in the specification and claims are to be understood as being modified in all instances by the term “about.” Accordingly, unless indicated to the contrary, the numerical parameters set forth in this specification and attached claims may be applied in the present invention as approximations that may vary depending upon the desired properties sought to be obtained by embodiments of the present disclosure. As used herein, “about” may be understood by persons of ordinary skill in the art and can vary to some extent depending upon the context in which it is used. If there are uses of the term which are not clear to persons of ordinary skill in the art, given the context in which it is used, “about” may mean up to plus or minus 10% of the particular term. In another embodiment the recited numerical parameter is employed instead of its approximation.

The word “comprise” and variations thereof such as “comprises” and “comprising” are meant inclusively and include additional possible components that are technically compatible as understood by a person skilled in the art. These terms also may in certain embodiments include their narrow meaning “consisting of”.

The generally accepted IUPAC single letter abbreviations for amino acids and their side chains in polypeptides are used herein. In accordance with 20 standard nomenclature, amino acid residue sequences are denominated by either a three letter or a single letter code as indicated as follows: Alanine (Ala, A), Arginine (Arg, R), Asparagine (Asn, N), Aspartic Acid (Asp, D), Cysteine (Cys, C), Glutamine (Gln, Q), Glutamic Acid (Glu, E), Glycine (Gly, G), Histidine (His, H), Isoleucine (Ile, I), Leucine (Leu, L), Lysine (Lys, K), Methionine (Met, M), Phenylalanine (Phe, F), Proline (Pro, 25 P), Serine (Ser, S), Threonine (Thr, T), Tryptophan (Trp, W), Tyrosine (Tyr, Y), and Valine (Val, V). For example, Cys^1 means amino acid cysteine in the position 1 (first amino acid) of the amino acid sequence, such as a peptide.

The term "variant", including functional variants, denotes a polypeptide or polynucleotide that deviates from a reference polypeptide or polynucleotide while maintaining fundamental characteristics. An exemplary variant of a polypeptide varies in its amino acid sequence compared to another reference polypeptide. Generally, these disparities are restricted so that the sequences of the reference polypeptide and the variant are closely similar overall, with substantial identity observed in numerous regions. Discrepancies in amino acid sequence between a variant and reference polypeptide may arise from one or more modifications, including substitutions, additions, and/or deletions. A substituted or inserted amino acid residue may or may not be encoded by the genetic code.

In an embodiment the functional variant has similar binding affinity and/or binding specificity, as the homing peptide. In another embodiment this functional similarity is at least 80%, at least 90%, at least 95%, or at least 98% of that of the homing peptide.

A variant of a polypeptide may be either naturally occurring, such as an allelic variant, or as a variant not documented in natural contexts. Structural modifications and alterations can be introduced to the polypeptide described herein while still yielding a molecule possessing comparable characteristics to the original polypeptide (e.g., through conservative amino acid substitutions). For instance, specific amino acids within a sequence can be interchanged with others without significant loss of functionality/activity. Given that the biological functional activity of a polypeptide is chiefly determined by its interactive capacity and nature, certain substitutions in the amino acid sequence can be implemented without compromising the essential properties of the polypeptide, thereby resulting in a polypeptide with similar attributes.

Modifications and alterations in the structure of the disclosed polypeptides can be implemented to generate a molecule with similar characteristics to the original polypeptide, for example a conservative amino acid substitution or conservative amino acid substitutions. Notably, specific amino acids within a sequence can be exchanged without significant loss of activity. The interactive capacity and nature of a polypeptide are pivotal in defining its biological functional activity, allowing for certain amino acid sequence substitutions without compromising the essential properties, resulting in a polypeptide with comparable attributes.

In effecting such changes, the hydropathic index of amino acids becomes a relevant consideration. The significance of the hydropathic amino acid index in conferring interactive biological function on a polypeptide is widely acknowledged in the field. It is recognized that substituting certain amino acids with others having a similar hydropathic index can yield a polypeptide with similar biological activity. Each amino acid is assigned a hydropathic index based on its hydrophobicity and charge characteristics. Those indices are: isoleucine (+4.5); valine (+4.2); leucine (+3.8); phenylalanine (+2.8); cysteine/cysteine (+2.5); methionine (+1.9); alanine (+1.8); glycine (-0.4); threonine (-0.7); serine (-0.8); tryptophan (-0.9); tyrosine (-1.3); proline (-1.6); histidine (-3.2); glutamate (-3.5); glutamine (-3.5); aspartate (-3.5); asparagine (-3.5); lysine (-3.9); and arginine (-4.5). The relative hydropathic character of the amino acid is believed to influence the secondary structure of the resultant polypeptide, defining its interaction with other molecules, for example enzymes, substrates, receptors, antibodies, antigens, and the like.

Amino acid substitutions can be made on the basis of the hydropathic index, with preferred substitutions falling within ± 2, particularly preferred within ± 1, and even more particularly preferred within ± 0.5. Similar substitutions based on hydrophilicity values are also feasible, especially in cases where the resultant polypeptide or peptide is intended for immunological applications. Hydrophilicity values assigned to amino acid residues guide such substitutions, with preferred values falling within ± 2, particularly preferred within ± 1, and even more particularly preferred within ± 0.5. The following hydrophilicity values have been assigned to amino acid residues: arginine (+3.0); lysine (+3.0); aspartate (+3.0 ± 1); glutamate (+3.0 ± 1); serine (+0.3); asparagine (+0.2); glutamine (+0.2); glycine (0); proline (-0.5 ± 1); threonine (-0.4); alanine (-0.5); histidine (-0.5); cysteine (-1.0);methionine (- 1.3); valine (-1.5); leucine (-1.8); isoleucine (-1.8); tyrosine (-2.3); phenylalanine (-2.5); tryptophan (-3.4).

It is understood that amino acids can be substituted based on hydrophilicity values, with preferred substitutions within ± 2, particularly preferred within ± 1, and even more particularly preferred within ± 0.5, to obtain a biologically equivalent, and specifically an immunologically equivalent polypeptide. Amino acid substitutions, as outlined, are generally informed by the relative similarity of side-chain substituents, for example hydrophobicity, hydrophilicity, charge, size, and other characteristics. Various exemplary substitutions, considering these characteristics, include (original residue: exemplary substitution): (Ala: Gly, Ser), (Arg: Lys), (Asn: Gln,His), (Asp: Glu, Cys, Ser), (Gln: Asn), (Glu: Asp), (Gly: Ala), (His: Asn, Gln), (Ile: Leu,Val), (Leu: Ile, Val), (Lys: Arg), (Met: Leu, Tyr), (Ser: Thr), (Thr: Ser), (Tip: Tyr), (Tyr:Trp, Phe), and (Val: Ile, Leu). This disclosure contemplates functional or biological equivalents of the polypeptide, including variants with approximately 50%, 60%, 70%, 80%, 90%, and 95% sequence identity to the polypeptide of interest.

In an embodiment the homing peptide or its functional variant has C as the N-terminal residue. Embodiments having C as the N-terminal residue are advantageous because they allow for example chemical modification by acylation-based radiolabelling, and/or attachment of another chemical moiety.

In an embodiment the homing peptide or its functional variant has C as the N-terminal residue and optionally A as the C-terminal residue.

In an embodiment the homing peptide or its functional variant has CG as the N-terminal residues.

In an embodiment the homing peptide or its functional variant has CGR as the N-terminal residues.

In an embodiment the homing peptide or its functional variant has A as the C-terminal residue.

In an embodiment the homing peptide or its functional variant has VA as the C-terminal residues.

In an embodiment the homing peptide or its functional variant has RVA as the C-terminal residues.

In an embodiment the homing peptide comprises at least six amino acids. In an embodiment the homing peptide comprises 6-30 amino acids, 6-20 amino acids, 6-15 amino acids, or 6-10 amino acids.

In an embodiment the homing peptide consists of at least six amino acids. In an embodiment the homing peptide consists of 6-30 amino acids, 6-20 amino acids, 6-15 amino acids, or 6-10 amino acids.

In an embodiment the homing peptide comprises or consists of at 6-15 amino acids.

In an embodiment the homing peptide comprises at least nine amino acids. In an embodiment the homing peptide comprises 9-30 amino acids, 9-20 amino acids, 9-15 amino acids, or 9-10 amino acids.

In an embodiment the homing peptide consists of at least nine amino acids. In an embodiment the homing peptide consists of 9-30 amino acids, 9-20 amino acids, 9-15 amino acids, or 9-10 amino acids.

In an embodiment the homing peptide specifically binds to a metastatic brain cell or other cancerous cell. Preferably the homing peptide specifically binds to a target protein in the metastatic brain cell. The homing peptide can simultaneously treat and/or diagnose a disease or a condition at one or more locations in the body.

In an embodiment the homing peptide specifically binds to PRDX1, peroxiredoxin 1. The capability of the homing peptide, or a variant thereof, to bind specifically to PRDX1 can be analysed by the method described in the present description, such as in Example 6 or 7.

The expression “specifically binds” means that the homing peptide is capable of selectively attaching and/or adhering to a specific target molecule, typically a biomolecule such as a protein. The term specifically binds further implies that the homing peptide has a high affinity for its intended target, preferably PRDX1. In an embodiment the affinity is higher than that of a control peptide. When specifically binding, the homing peptide forms a selective and strong, but reversible, interaction with a particular target molecule, allowing for accurate

detection, quantification, or tracking of the target. In another embodiment the interaction, i.e. the specific binding to the target molecule is irreversible.

In an embodiment the RVA binds to PRDX1 with micromolar binding affinity.

According to another aspect is provided a complex comprising the homing peptide and a target molecule, such as an RVA-PRDX1 complex.

In an embodiment the target molecule has 70%, 75%, 80%, 85%, 90%, 95%, or 99% sequence identity to PRDX1, or to its N-terminal domain.

In an embodiment the target molecule is PRDX1, or its N-terminal domain.

According to another aspect is provided a complex composition comprising the present complex. In an embodiment the complex composition comprises at least one of buffering agent, preservative, and water.

In another embodiment is provided a complex comprising the homing peptide bound non-covalently to PRDX1 or to its N-terminal domain.

In an embodiment the homing peptide is internalized by a metastatic brain cell or other cancerous cell. Internalization of the homing peptide by a metastatic brain cell or other cancerous cell can be analysed and verified by the method of Example 2.

In an embodiment the homing peptide does not home significantly to liver, lungs, kidneys, or spleen.

In an embodiment the homing peptide may be chemically modified. Preferably the chemical modification is in one or more amino acid of the homing peptide, such as in an amino acid in position 1,2,3,4,5,6,7,8, or 9 of the homing peptide or in the linker molecule.

In an embodiment at least one amino acid of the homing peptide is acetylated.

In an embodiment the homing peptide is a cyclic peptide, or it is a part of a cyclized peptide. Preferably the cyclic peptide is formed by cyclizing two amino acids in the amino acid sequence of the homing peptide. Alternatively, one amino acid of the homing peptide participates in the formation of the cyclic peptide, the other amino acid being present outside the amino acid sequence of any one of SEQ ID NO: 1-6.

In an embodiment the homing peptide contains one or more unnatural amino acid. Preferably the one or more unnatural amino acid is present in the amino acid sequence of the homing peptide in a position marked with X in the amino acid sequences disclosed herein. Alternatively or additionally, the unnatural amino acid may be present in positions outside the sequence specified by any of the sequences SEQ ID NO: 1-6.

In an embodiment the unnatural amino acid is one or more individually selected from: D-amino acids, homo-amino acids, methylated amino acids, beta-homo-amino acids, N-methyl amino acids, alpha-methyl amino acids, citrulline (Cit), hydroxyproline (Hyp), norleucine (Nle), 3-nitrotyrosine, nitroarginine, ornithine (Orn), naphtylalanine (Nal), aminobutyric acid (Abu), 2,4-diaminobutyric acid (DAB), p-Benzoyl-L-phenylalanine (Bpa), p-Azido-tetrafluoro-L-phenylalanine (ZpN3), methionine sulfoxide, and methionine sulfone.

In an embodiment the homing peptide is attached to a photoreactive unnatural amino acid (Bpa or ZpN3) for the covalent conjugation to the target protein.

In an embodiment the homing peptide is attached to biotin, in addition to the photoreactive unnatural amino acid, for the detection and pull down of the target protein(s). The covalent crosslinking of the homing peptide to the target protein(s) is described in methods of example 3.

In an embodiment the C-terminus of the homing peptide is amidated.

In an embodiment the N-terminus of the homing peptide is not chemically modified. According to this embodiment, the N-terminal amino acid has an N-terminal NH_2 group.

In an embodiment the C-terminus of the homing peptide is not chemically modified. According to this embodiment, the C-terminal amino acid has a C-terminal COOH group.

In an embodiment the N-terminus and/or the C-terminus of the homing peptide is acylated, preferably acetylated.

In an embodiment in the conjugate the cargo molecule is covalently or non-covalently linked, directly or indirectly conjugated to or complexed with the N-terminal end of the homing peptide.

In an embodiment in the conjugate the cargo molecule is covalently or non-covalently linked, directly or indirectly conjugated to or complexed with the C-terminal end of the homing peptide.

In an embodiment in the conjugate the homing peptide is non-covalently bound to the cargo molecule.

In an embodiment the conjugate contains an unnatural amino acid between the homing peptide and the cargo molecule.

In an embodiment the homing peptide is chemically modified by a modification disclosed in Table 3.

In an embodiment in the conjugate the cargo molecule is an imaging agent, preferably a fluorescent dye, more preferably TAMRA (carboxytetramethylrhodamine), which is an amine-reactive fluorescence molecule used for labelling oligonucleotides and/or proteins and/or peptides.

In another embodiment in the conjugate the cargo molecule is a target for pull-down, preferably Biotin.

In an embodiment in the conjugate the cargo molecule is a drug molecule. The drug molecule can be a cytotoxic agent, antineoplastic agent,$_D$(KLAKLAK)$_2$, SEQ ID NO: 7, or MMAE. In another embodiment the cargo molecule is at least one of nucleic acid payload, antibody cargo, monoclonal antibody cargo, tagged molecule, fluorescent tag molecule, nanoparticle cargo, gene therapy cargo, RNA payload, liposomal payload, siRNA cargo, conjugated nanoparticle, peptide moiety, and protein-based cargo.

In another embodiment in the conjugate the cargo molecule is one or more of a nucleic acid (example agents can include antisense molecules, aptamers, ribozymes, triplex forming oligonucleotides, external guide sequences, RNAi, CRISPR/Cas, zinc finger nucleases, and transcription activator-like effector nucleases); the nucleic acid can be an expression vector encoding a protein or a functional nucleic acid, vectors can be suitable for integration into a cells genome or expressed extra-chromosomally, antibody or monoclonal antibody, Tagged molecule, Fluorescence tag molecule, label, nanoparticle, Conjugated nanoparticle, Liposomal payload, small molecule (example: anti-cancer agent, an anti-inflammatory agent, an immunomodulatory agent, and an antimicrobial agent); include organic and organometallic compounds; can be hydrophilic, hydrophobic, or amphiphilic compounds, toxic agent, radioisotope, dye, a radionuclide, a fluorescent tag/compound/dye, a magnetic tag/molecules, paramagnetic molecules, x-ray imaging agents, and contrast media, dyes, near infra-red dyes, SPECT imaging agents, PET imaging agents, peptide, lipid, glycolipid, glycoprotein, polypeptide or broader: therapeutic agent, a diagnostic agent, a prophylactic agent, a nutraceutical agent.

In an embodiment the present homing peptide, its functional variant, or a conjugate comprising such, does not comprise a sequence DACSRFLGERVDATAAGCSR, SEQ ID NO: 56.

In an embodiment the present homing peptide, its functional variant, or a conjugate comprising such, does not comprise a sequence CARRLGRVATTYYMDVW, SEQ ID NO: 57.

In an embodiment the present homing peptide, its functional variant, or a conjugate comprising such, does not comprise a sequence CARRLGRVARPTTWTS, SEQ ID NO: 58.

In an embodiment the present homing peptide, its functional variant, or a conjugate comprising such, does not comprise a sequence CGVRLGC, SEQ ID NO: 59.

In an embodiment the present homing peptide, its functional variant, or a conjugate comprising such, does not comprise a sequence CVPELGHEC, SEQ ID NO: 60.

The following numbered embodiments 1-18 are disclosed:

1. A homing peptide, or its functional variant, comprising an amino acid sequence CxxxLG.
2. The homing peptide or its functional variant according to embodiment 1 comprising an amino acid sequence CxxxLGxxx.
3. The homing peptide or its functional variant according to embodiment 1 comprising an amino acid sequence CxRxLGRVA, SEQ ID NO: 3.
4. The homing peptide or its functional variant according to embodiment 1 or 2 comprising an amino acid sequence CGRTLGRVA, SEQ ID NO: 4.
5. The homing peptide or its functional variant according to any one of embodiments 1-4, wherein the homing peptide has no more than 30 amino acids.
6. The homing peptide or its functional variant according to any one of the embodiments 1-5 capable of specifically binging PRDX1 (peroxiredoxin 1), preferably an N-terminal domain of PRDX1.
7. A conjugate comprising the homing peptide or its functional variant of any one of the embodiments 1-6, and a cargo molecule, wherein the cargo molecule is at least one of: a nucleic acid, antisense molecule, aptamer, ribozyme, triplex forming oligonucleotide, external guide sequence, RNAi, CRISPR/Cas, zinc finger nuclease, transcription activator-like effector nucleases, an expression vector encoding a protein or a functional nucleic acid, a vectors suitable for integration into a cells genome or expressed extra-chromosomally, antibody, monoclonal antibody, tagged molecule, fluorescence tag molecule, label, nanoparticle, conjugated nanoparticle, liposomal payload, small molecule, anti-cancer agent, an anti-inflammatory agent, an immunomodulatory agent, an antimicrobial agent, an organic compound, an organometallic compound; a hydrophilic compound, a hydrophobic compound, an amphiphilic compounds, toxic agent, radioisotope, a radionuclide, a fluorescent tag/compound/dye, a magnetic tag/molecule, paramagnetic molecule, x-ray imaging agent, contrast media, dye, near

infra-red dyes, SPECT imaging agent, PET imaging agent, peptide, lipid, glycolipid, glycoprotein, polypeptide, a therapeutic agent, a diagnostic agent, a prophylactic agent, a nutraceutical agent.

8. The conjugate of embodiment 7, wherein the cargo molecule is an imaging moiety, or a diagnostic moiety, alone or in combination with one or more therapeutic moiety.

9. The conjugate of embodiment 7, wherein cargo molecule is an imaging agent.

10. A pharmaceutical composition comprising the conjugate of any one of the embodiments 7-9 and a pharmaceutically acceptable carrier.

11. The homing compound of any one of embodiments 1-6, or the conjugate of any one of embodiments 7-9, for use as a medicine.

12. The homing compound of any one of embodiments 1-6, or the conjugate of any one of embodiments 7-9, for treating cancer, preferably a brain metastasis.

13. A method of *in vivo*, or *ex vivo*, or *in vitro* diagnosis of cancer using the homing compound of any one of embodiments 1-6, or the conjugate of any one of embodiments 7-9, wherein the cancer is preferably at least one of melanoma, breast carcinoma, lung carcinoma, brain cancer, brain metastatic cell, or any other cancer with PRDX1 expression.

14. A method of delivering a cargo molecule into a cancer cell comprising contacting the cancer cell with the conjugate of any one of embodiments 7-9.

15. A method of identifying a cancerous target cell or cancerous target tissue in a biological sample, the method comprising:

(a) contacting the biological sample with the conjugate of any one of embodiments 7-9 for a time and under conditions sufficient for specifically binding the conjugate to the target cell or the target tissue; and

(b) detecting the presence or absence of the bound conjugate in the biological sample;

wherein presence of the conjugate in the detecting step (b) indicates that the cancerous target cell or the cancerous target tissue is present in the biological sample.

16. The method of embodiment 15, wherein the step (b) is carried out by immunohistochemistry, fluorescent imaging, radioimmunoassay, or immunofluorescence.

17. A kit comprising: (a) a container comprising a pharmaceutical composition containing the homing peptide of any one of embodiments 1-6, the conjugate of any one of

embodiments 7-9, or the pharmaceutical composition of embodiment 10, in solution or in a lyophilized form; (b) optionally, a second container containing a diluent or reconstituting solution for the lyophilized formulation; and (c) optionally, instructions for use of the solution or reconstitution and/or use of the lyophilized formulation.

18. A complex comprising the homing peptide of any one of the embodiments 1-6 and a target molecule.

EXAMPLES

The following examples are provided to better illustrate the claimed invention. The examples are not to be interpreted as limiting the scope of the invention, which is determined by the claims. Any example or embodiment described in this application and not falling within the scope of the independent claims is presented herein as an additional example or additional embodiment useful for understanding the claimed invention. To the extent that specific materials are mentioned, it is merely for purposes of illustration and is not intended to limit the invention. One skilled in the art may develop or purchase equivalent means or products to carry out the claimed invention without exercising inventive capacity and without departing from the scope of the invention. It shall be understood that many variations can be made in the procedures described herein while remaining within the scope of the present invention.

Example 1 – Identification of novel brain metastases homing peptide

We engineered the fatty acid binding protein 3 (FABP3) overexpressing MDA-MB-231 triple negative breast cancer (TNBC) cell line (here after named as the invasive variant i231) along with the parental MDA-MB-231 cell line (231) as well as another TNBC cell line, MDA-MB-468 (468) to express the green fluorescence protein (GFP) and luciferase and inoculated them through the heart into the immunocompromised NMRI nu/nu female mice. Tumour growth was monitored by using the whole-body bioluminescence imaging (BLI) following intraperitoneal injection of luciferin. After detection of tumour formation in the brain, the brains were excised and dissected into two halves. One half was processed for cell culture and the second half for histology. The brain seeking cells (i231, 231 or 468) were isolated, expanded in the culture, and injected through the heart for subsequent selection cycles (Fig. 1A). FAPB3 expression promoted brain metastasis formation as we detected tumour cells in the brains of 40% (4/10) of the i231-injected mice after the first intracardiac injection compared to 10% (1/10) of 231-injected mice and 12,5% (1/8) of 468-injected mice (Fig. 1B). The i231 cells also grew faster in the brain as the brain metastasis formation was detected by histological analysis and the isolated tumour cells expanded in the culture. On

the other hand, no 231-cell derived tumour formation was detected in histological analysis but the cells from one mouse grew in culture while with the 468 derived tumour cells were detected in the brain only with the BLI and no tumours were detected in the histological analysis and no tumour cells enriched during the cell culture (Fig. 1B). After the second selection round brain metastasis formation was increased as 50% (5/10) of the i231-injected and 50% (2/4) of the 231-injected mice formed brain metastasis (Fig. 1C). In addition, the time for the brain metastasis formation was significantly faster (1-2 months) than after the first selection round indicating successful selection of brain seeking cells. Fig. 1D-E show the representative histology of tumour cells in the mouse brain after first and second selection rounds of the established brain metastasis.

To identify peptides that would specifically target brain metastases, we setup a phage displayed peptide screening pipeline using brain metastatic cells derived from the i231 cells after the second selection round (i231-BrM2) (Fig. 1F-H). After two rounds of *ex vivo* and three rounds of *in vivo* selection using two different phage-displayed peptide libraries (CX_7C and CX_8C; C = cysteine, X = any amino acid residue), five different enriched peptides (LTC, RVA, TRG, ART, VSN) were identified and selected based on sequencing the part of the phage genome that encodes the peptides (Table 1, Fig. 1F-J). These five individual phages were further validated using *ex vivo* binding of each phage to the i231-BrM2 tumour containing brain cell suspension. The phage displaying the G7 (seven glycine residues) peptide was used as a control peptide to determine the background binding. Only the RVA and VSN phage showed better binding than the control phage (Fig. 1K). Next, we injected the control, RVA, ART, and VSN phage into the tail vein of i231-BrM2 brain tumour bearing or healthy mice to assess the *in vivo* homing to brain metastatic tumours. The RVA displaying phage showed superior and specific homing to the BrMs compared to other phage and the healthy brain (Fig. 1L).

Table 1: Enriched single phage clones identified using the ex vivo/in vivo phage display.

Library	Peptide sequence	Short sequence	Length	In vivo R1	In vivo R2	Total no. of occurrence
CX_7C	CAGTECLTC	LTC	9	10		10
	CGRTLGRVA	RVA	9		2	2
	CGRTRG	TRG	6	9	17	26
	CSHVSN	VSN	6	5		5
CX_8C	CKLLSART	ART	8	17		17

Example 2 – RVA internalizes in the BrM2 cells and homes to the BrM2 derived intracranial tumours

To validate that the homing of RVA to brain metastases was mediated by the RVA peptide sequence, a synthetic RVA or a control (Ctrl) peptide (CVAALNADG, SEQ ID NO: 8) was conjugated to the TAMRA (TMR) fluorescence dye. First, we studied the uptake of the TMR-RVA or TMR-Ctrl peptide (0.1 µg/mL) by the i231-BrM2 cells after 60 minutes incubation at 37°C using microscopic evaluation. TMR-RVA was very efficiently taken up by the i231-BrM2 cells while no uptake was detected in cells incubated with the TMR-Ctrl peptide (Fig. 2A, 2B). Next, we injected the TMR-RVA or TMR-Ctrl peptide intravenously into the tail vein of mice bearing intracranial i231-BrM2 tumours engineered to express GFP. After 30 or 60 minutes of circulation, mice were sacrificed, and the brain and control organs (liver, lungs, kidneys, and spleen) were excised and prepared for immunohistochemistry. TMR-RVA showed significantly stronger signal in the GFP-positive tumours compared to TMR-Ctrl peptide indicating specific homing to the metastatic cells (Fig. 2C, Fig. 9A). RVA homing also increased in a time dependent manner as judged by quantification of the fluorescence signal in the tumours at 30 and 60 minutes after injection (Fig. 2D). To evaluate the specificity of RVA homing to the i231-BrM2 tumour cells, we studied the distribution of RVA in the peripheral organs relative to the TMR-Ctrl peptide. Negligible amount of RVA was detected in the liver, lungs, kidneys, or spleen 60 minutes after intravenous injection indicating very specific homing to the metastatic brain tumours (Fig. 2E-F). Interestingly, no TMR-RVA homing was detected to the orthotopic mammary fat pad implanted MDA-MB-231 tumours (Fig. 9B).

Example 3 - Photoaffinity labelling reveals the RVA target protein

As RVA so specifically targets brain metastases *in vivo*, it must recognize a BrM-specific target. To identify the target, we injected the i231-BrM2-bearing mice systemically with 100 µM biotinylated RVA or biotinylated control peptide and allowed them to circulate for 60 minutes. Brain tissue extracts were prepared for Western Blot analysis by probing the membrane with Streptavidin-HRP (SA-HRP) to detect the biotinylated peptides with their bound target proteins. Western Blot analysis revealed an approximately 22kDa band when probed with the biotinylated RVA (Fig. 3A Lane 1) while no band was detected when probed with the biotinylated control peptide (Fig. 3A, Lane 2). Next, we designed two different photoreactive crosslinker variants of the biotinylated RVA by incorporating an unnatural amino acid, Bpa or ZpN3 (Fig. 3B, C) to the N-terminus of RVA to create biotinylated-A-Bpa-RVA (Bpa-RVA) and biotinylated-A-ZpN3-RVA (ZpN3-RVA). Bpa and ZpN3 crosslinking is activated at UV wavelengths of 365nm and 254 nm, respectively. First, we

incubated the i231-BrM2 cell extract with the Bpa-RVA or ZpN3-RVA followed by UV irradiation (UV Crosslinker Stratalinker 1800) to crosslink RVA with the target. Extracts were then subjected to Western Blot analysis and the membranes were probed with SA-HRP to detect the crosslinked protein. Also, this Western Blot analysis revealed a 22kDa protein band (Fig. 3D, E) only in the lanes that contained the crosslinked RVA peptide. Quantification showed that the intensity of photoaffinity labelling increased with the increased UV irradiation time from 15 min to 30 min (Fig. 3F, G).

For the pull down of the RVA target protein from the tumour tissue, we incubated the photoreactive RVA (Biotin-A-Bpa-RVA) for 60 minutes with the metastases containing brain extract and performed the UV irradiation for 30 min. The protein(s) crosslinked to RVA was then pulled down by using streptavidin beads. The RVA and crosslinked target was then eluted from the beads by addition of free biotin and visualized using silver staining (Fig. 3H). A 22kDa specific band was cut out from the gel (lane 3) and analysed by mass spectrometry (MS) to reveal the target protein candidates (Table 2). The identified proteins were named and classified using the UniProt Knowledgebase database (Apweiler et al. 2004).

Based on the MS data, peroxiredoxin 1 (PRDX1) was the only 22kDa protein detected and showed the highest sequence coverage (29%) of the pulled-down proteins. To validate the specificity of the RVA binding to the 22kDa protein, we performed a competition binding study utilizing i231-BrM2 cell extracts incubated overnight with either Bpa-RVA (Fig. 3I) or ZpN3-RVA (Fig. 3J) with increasing concentrations of naked original RVA. After 30 minutes of UV irradiation, extracts were probed with SA-HRP and analysed by Western Blot. The target protein band (22kDa) intensity decreased with increasing concentrations of the naked RVA (Fig. 3I, J). Quantification of the amount of crosslinked protein showed that already 1 µM of naked RVA decreased the crosslinking by approximately 50% and with 100 µM of naked RVA, the 22kDa band totally disappeared (Fig. 3K, 3L) indicating that RVA was able to compete for the binding of the biotinylated Bpa-RVA or ZpN3-RVA.

Table 2: RVA binding protein candidates identified by mass spectrometry.

Protein ID	Gene Name	Description	Protein functions	Coverage (%)	Molecular weight (kDa)
Q06830	PRDX1	Peroxiredoxin 1	Oxidation/Reduction	29	22
Q13162	PRDX4	Peroxiredoxin 4	Oxidation/Reduction	7	31
P60201	PLP1	Myelin proteolipid protein	CNS myelin protein	20	30
O60814	H2BC12	Histone H2B type 1-K	DNA binding protein	19	13.9
P22626	HNRNPA2B1	Heterogeneous nuclear ribonucleoproteins A2/B1	RNA binding protein	17	36
O95716	RAB3D	Ras-related protein Rab-3D	Vesicular trafficking	13	24
P07900	HSP90AA1	Heat shock protein HSP90-alpha	Molecular chaperone	10	84.7
P0DMV8	HSPA1A	Heat shock 70kDa protein 1A	Molecular chaperone	5	70
P11142	HSPA8	Heat shock cognate 71kDa protein	Molecular chaperone	5	70.9
A0A075B759	PPIAL4E	Peptidyl-prolyl cis-trans isomerase A-like 4E	Molecular chaperone	8	59
P68104	EEF1A1	Elongation factor 1-alpha 1	Protein biosynthesis	5	34
P60842	EIF4A1	Eukaryotic initiation factor 4A-1	Protein biosynthesis	9	46
P07195	LDHB	L-lactate dehydrogenase B chain	Glycolysis	6	37
P60174	TPI1	Triose phosphate isomerase	Glycolysis	10	26
P06733	ENO1	Alpha-enolase	Glycolysis	16	47.1
P04406	GAPDH	Glyceraldehyde-3-phosphate dehydrogenase	Housekeeping/Glycolysis	20	37
P08670	VIM	Vimentin	Cytoskeletal protein	8	53.7
P60709	ACTB	Actin, cytoplasmic 1	Cytoskeletal protein	19	40.4
P0DPH7	TUBA3C	Tubulin alpha-3C chain	Cytoskeletal protein	3	50
P07737	PFN1	Profilin-1	Cytoskeletal protein	21	15

Example 4 - RVA targets PRDX1 *in vitro* and *in vivo*

To further validate the RVA binding to the PRDX1, we silenced PRDX1 in the i231-BrM2 cells using three different shRNAs (sh1, sh2 and sh3). Silencing efficiency was confirmed by Western Blot analysis using anti-PRDX1 antibodies (Fig. 4A). PRDX1 protein expression level was quantified relative to the loading control, β-tubulin (Fig. 4A) and the mRNA expression was quantified by qPCR (Fig. 4B). All shRNAs silenced the protein level expression very efficiently and mRNA expression was decreased by 67%, 73% and 93% with sh1, sh2 and sh3, respectively. When we incubated the i231-BrM2 cells transduced with lentiviruses encoding the non-targeting shCtrl with 100 ng/mL TMR-RVA for 60 minutes followed by staining with anti-PRDX1 antibodies, strong colocalization between PRDX1 (white) and RVA (red) inside the cells was detected (Fig. 4C, upper panel). On the contrary, PRDX1-silenced cells showed no RVA internalization nor PRDX1 expression (Fig. 4C, lower panel) indicating that PRDX1 is required for RVA internalization. To further validate the interaction between PRDX1 and RVA, we performed the photoactivable crosslinking using non-silenced control (shCtrl) or PRDX1-silenced i231-BrM2 cell extracts and Bpa-RVA or ZpN3-RVA. After UV irradiation, the Western blot analysis did not detect any protein

band at 22kDa in the PRDX1 silenced cells (sh2 and sh3) while it was clearly visible in the control cells (Fig. 4D, E) confirming the crosslinking between RVA and PRDX1.

To study whether PRDX1 is expressed by the brain metastatic cell lines, we performed Western Blot analysis of the cell extracts of breast cancer brain metastatic cell lines (231-BrM2, i231-BrM2) and their corresponding primary cell lines (231 and i231). Very similar PRDX1 levels were detected in all tested cell lines (Fig. 4F). We silenced PRDX1 in the i231-BrM2 cells (Fig. 4G) and investigated whether PRDX1 would be required for cell proliferation or viability. Starting at day 10 post-silencing we followed the cell proliferation for four days by using the MTT assay. A significant decrease in cell proliferation/viability in PRDX1-silenced i231-BrM2 cells compared to non-silenced control cells was detected (Fig. 4H) suggesting that PRDX1 is required for the proliferation/viability of these cells.

The PRDX1 expression was investigated in melanoma derived brain metastatic cell lines and their corresponding primary melanoma cell lines. Majority (4/6) of the tested cell lines expressed low amounts of PRDX1 and two cell lines (RALL.CB2 and RALL.C) showed high PRDX1 expression (Fig. 4I). We silenced PRDX1 in RALL.CB2 cells to study the effect on cell proliferation/viability (Fig. 4J). Cell proliferation of the RALL.CB2 cells was followed for four days starting five days post-silencing. Significant decrease in cell proliferation/viability was seen in PRDX1-silenced cells compared to non-silenced control cells (Fig. 4K).

Homing of the TMR-RVA to the i231-BrM2 derived intracranial tumours was studied 60 min after intravenous injection. TMR-RVA (red in original image) accumulated in the tumour (green) area that expressed PRDX1 (white in orig. image) (Fig. 4L). No RVA (red in original image) homing was detected in the surrounding healthy tissue. The crosstalk between the metastatic lesion (green in orig. image) and the surrounding microenvironment was also investigated (Fig. 10A-10B and 10a-10d). In addition to the PRDX1, the reactive astrocytes (GFAP), resident microglia (IBA1), neurons (TUJ1), showed activated expressions in the tumour area compared to the surrounding healthy tissue (Fig. 10a-d).

Example 5 – Cys^{1}, Leu^{5}, Gly^{6} in the RVA peptide are required for its cellular uptake

We studied the sequence-activity relationships (SAR) of the RVA variants by *in vitro* competition binding assay of the TMR-RVA in the absence or presence of its non-fluorescent variants. Since RVA showed cell penetrating properties, we first studied which amino acids in the RVA sequence would be required for its uptake by the brain metastatic cells. We substituted each amino acid in the RVA sequence with Alanine in a sequential order to generate eight peptide variants [Ala^{1}]RVA – [Ala^{8}]RVA. Further modifications to the RVA sequence including a retroinverse variant or introduction of Lysine (K) residue to the

C-terminus (RVAK) were also synthesized. In addition, the free native carboxylic acid (COOH) group at the C-terminus was amidated or the N-terminus acetylated or both. Furthermore, new RVA variants were generated by replacing each amino acid residue with their D-isomer counterparts (Table 3).

Table 3: RVA and its variants used in this study.

Abbreviation	Sequence	Purity	Molecular weight (g/mol)
RVA	H-CGRTLGRVA-NH_2	98	1273.20
	Ac-CGRTLGRVA-NH_2	99	1201.20
	H-CGRTLGRVA-OH	98	1274.20
	Ac-CGRTLGRVA-OH	99	1202.20
[Ala^1] RVA	H-AGRTLGRVA-NH_2	99	1241.14
[Ala^2] RVA	H-CARTLGRVA-NH_2	96	1287.23
[Ala^3] RVA	H-CGATLGRVA-NH_2	95	1074.07
[Ala^4] RVA	H-CGRALGRVA-NH_2	99	1243.17
[Ala^5] RVA	H-CGRTAGRVA-NH_2	99	1231.12
[Ala^6] RVA	H-CGRTLARVA-NH_2	95	1287.23
[Ala^7] RVA	H-CGRTLGAVA-NH_2	97	1074.07
[Ala^8] RVA	H-CGRTLGRAA-NH_2	98	1245.14
Retroinverse-RVA	Ac-avrglrtgc-NH_2	98	1201.21
[D-Cys^1] RVA	H-cGRTLGRVA-NH_2	99	1273.20
[D-Arg^2] RVA	H-CGrTLGRVA-NH_2	99	1273.20
[D-Thr^4] RVA	H-CGRtLGRVA-NH_2	99	1273.20
[D-Leu^5] RVA	H-CGRTlGRVA-NH_2	99	1273.20
[D-Arg^7] RVA	H-CGRTLGrVA-NH_2	99	1273.20
[D-Val^8] RVA	H-CGRTLGRvA-NH_2	99	1273.20
[D-Ala^9] RVA	H-CGRTLGRVa-NH_2	99	1273.20
Cyclic 'native' RVA	H-CGRTLGRVC-NH_2	99	1303.24
RVAK	Ac-CGRTLGRVAK-NH_2	99	1443.41
Negative control	H-CVAALNADG-NH_2	97	945.97
TAMRA-A-RVA	TAMRA-ACGRTLGRVA-NH_2	99	1644.72
TAMRA-RVA	TAMRA-CGRTLGRVA-NH_2	99	1573.64
TAMRA-Negative control	TAMRA-CVAALNADG-NH_2	98	1245.40
Biotin-Bpa-A-RVA	Biotin-Bpa-ACGRTLGRVA-NH_2	99	1707.83
Biotin-ZpN3-A-RVA	Biotin-ZpN3-ACGRTLGRVA-NH_2	98	1716.70
Biotin-RVA	Biotin-CGRTLGRVA-OH	Commercial source	
Biotin-Negative control	Biotin-CVAALNADG-OH	Commercial source	
TrA Negative control	Ac-AGRTAARVA-NH_2	99	1141.10
KLAK-RVA	Ac-d(KLAKLAK)$_2$-GG-CGRTLGRVA-NH_2	98	3505.46
KLAK-TrA Negative control	Ac-d(KLAKLAK)$_2$-GG-AGRTAARVA-NH_2	99	3445.35
Free KLAK	d(KLAKLAK)$_2$	Commercial source	
MMAE-RVA	MMAE-CGRTLGRVA-NH_2	99	2630.32
MMAE-TrA Negative control	MMAE-AGRTAARVA-NH_2	99	2570.82
Free MMAE	MMAE	99	717.98

Underlined residues indicate an amino acid residue modification.

A competition study between each newly generated RVA variant and the TMR-RVA was performed using the i231-BrM2 cells. After 60 minutes of incubation with the TMR-RVA in the absence or presence of the non-labelled variants at varying concentrations, the TMR intensity-based competition binding affinity (IC_{50}) for each variant was measured (Fig. 11A-11R). The higher the IC_{50} value, the lower the cellular uptake activity and vice-versa. The [Ala^1]RVA (Fig. 11A) and [Ala^6]RVA Fig. 11F) variants with the highest IC_{50} values indicate the lowest cell penetrating activity while the [Ala^5]RVA (Fig. 11E) showed no internalization. Hence, Cysteine, Leucine and Glycine residues at positions 1, 5, and 6 (Cys^1, Leu^5, Gly^6) are required for the peptide's ability to compete with the TMR-RVA for the uptake by the i231-BrM2 cells. Alanine in any other position in the RVA sequence did not affect the uptake of the TMR-RVA by the i231-BrM2 cells indicating that amino acids in positions 2, 3, 4, 7, 8, and 9 do not contribute to the cellular uptake (Fig. 11B-11D, 11G-11H). In addition, N- or C-terminal modifications, addition of K in the C-terminus or retroinverso variant did not affect the internalization of the peptide (Fig. 11I-11P). Binding curves were determined for all the variants (Fig. 11Q). Based on this data, the three important amino acid residues (Cys^1, Leu^5, Gly^6) required for the RVA internalization were replaced with Alanine residues to create a new control peptide AGRTAARVA (abbreviated as TrA Ctrl, SEQ ID NO: 9) (Fig. 11R).

Example 6 - SPR and MST analyses confirm RVA peptide binding to the recombinant PRDX1

Next, we studied sequence-activity relationships (SAR) between the RVA variants and its recombinant target protein, PRDX1, using surface plasmon resonance (SPR) and microscale thermophoresis (MST). For the SPR binding study, 1.2 µg/mL N-terminally Histidine-tagged human recombinant PRDX1 protein (His-PRDX1) was affinity captured on the sample flow cells of a CM5 sensor chip (T100 Biacore machine) by 50 µg/mL of anti-Histidine antibody. As a reference control, an adjacent reference flow cell was coated with 50 µg/mL of anti-Histidine antibody but no His-PRDX1 was added. RVA or its variants were injected over the sensor chip surfaces at various concentrations (100 µM – 0.39 µM) for the binding analysis. A reference-subtracted sensorgram was generated and steady state binding affinity measured.

Similarly, MST was performed after labelling PRDX1 with a NT-650-NHS fluorescence dye. Sixteen different concentrations varying between 1000 µM and 0.03 µM as 2-fold serial dilutions of RVA or its variants were analysed for binding to 5 nM fluorescence PRDX1 using the Monolith NT.115 Pico instrument. Binding measurements were performed with the same settings including 20% excitation power, medium MST power, and 25°C temperature.

Both SPR and MST data were fitted using 1:1 binding model with the law of mass action to generate equilibrium dissociation constant (KD) as the readout for binding affinity. R2 denotes the goodness of the fit. Excellent fit is closest to 1 and the further away from 1 the worse the fit. In addition, signal-to-noise ratio (S/N) below five is considered poor quality MST data and/or no binding affinity while value above five is considered good quality data. RVA showed micromolar binding affinity to the recombinant PRDX1 as determined with both SPR and MST as well as high MST S/N (8.97) indicating an excellent binding. [Ala^2]RVA and [Ala^4]RVA showed micromolar binding affinity to PRDX1 in both SPR and MST data. The binding affinities of [Ala^1]RVA, [Ala^3]RVA, [Ala^5]RVA, and retroRVA to PRDX1 could not be calculated using SPR (> 1 mM) and no binding was detected using MST. [Ala^6]RVA, [Ala^7]RVA, [Ala^8]RVA, RVAK, control and TrA control showed no binding to PRDX1 in either SPR or MST. The SPR and MST data confirmed binding between PRDX1 and RVA and suggest that amino acid residues Cys^1, Arg^3, Leu^5, Gly^6, Arg^7, Val^8, and Ala^9 in the RVA sequence are important for the binding while the Gly^2 and Thr^4 are not required. Altogether, using the in vitro competition assay, we identified amino acid residues (Cys^1, Leu^5, Gly^6) required for the cellular uptake of the peptide and with SPR and MST analyses amino acid residue (Cys^1, Arg^3, Leu^5, Gly^6, Arg^7, Val^8, and Ala^9) required for the binding to recombinant PRDX1. This information will be useful for the design and modification of RVA peptide-based targeted applications. Results are shown in Table 4.

Table 4: SPR and MST binding affinities of His-humanPRDX1 with RVA or Alanine-derived variants.

RVA peptides	Binding Affinity, K_D (mean ± SD) µM	
Abbreviation	SPR	MST
RVA	184 ± 52	45 ± 17
[Ala^1] RVA	> 1000	ND
[Ala^2] RVA	61 ± 26	346 ± 49
[Ala^3] RVA	> 1000	ND
[Ala^4] RVA	56 ± 21	72 ± 33
[Ala^5] RVA	> 1000	ND
[Ala^6] RVA	ND	ND
[Ala^7] RVA	ND	ND
[Ala^8] RVA	ND	ND
RVAK	ND	ND
Retroinverse-RVA	> 1000	ND
Negative control	ND	ND
TrA Negative control	ND	ND

ND indicate no binding or not determined. K_D values are represented as mean ± SD.

Example 7 - Free N-terminus domain of PRDX1 (PRDX1-His) show improved binding affinity towards RVA and [D-Cys^1], [D-Leu^5] RVA variants.

As modification of the recombinant PRDX1 either via the N- or C-terminus may influence its binding to the RVA peptide. The crystal structures of PRDX1 in complex with various ligands, available in the Protein Data Bank (PDB ID codes: 7WET, 1QQ2, 2RII) showed that the two highly conserved cysteine residues at N-terminus (CYS52) and/or C-terminus (CYS173) are involved in ligand binding (Hirotsu et al. 1999; Jönsson, Johnson, and Lowther 2008; Xu et al. 2023). In addition, we are interested in the binding of RVA to the mouse variant of human PRDX1 since they share similar sequence identity with the human PRDX1.

To further characterize the RVA peptide, each amino acid residue in the RVA sequence was replaced with their D-isomer counterparts to generate new variants. The binding of RVA variants to C-terminally His-tagged recombinant human PRDX1 protein (PRDX1-His) was analysed by SPR. Our initial binding studies with N-terminally Histidine-tagged human recombinant PRDX1 (His-PRDX1) to RVA or its Alanine derived variants using both MST and SPR showed low (micromolar) binding affinities (Table 4).

Human PRDX1-His showed 2-fold higher binding affinity (KD = 86 ± 60 µM, Fig. 5A) to RVA compared to affinity obtained with the human His-PRDX1 (KD = 184 ± 52 µM, Table 4) thereby confirming that the free N-terminus of PRDX1 is involved in RVA binding. The [D-Cys1] affinity was very similar to PRDX1-His than that of the RVA (Fig. 5B), while the [D-Leu5] RVA variant showed almost 4-fold higher affinity to the human PRDX1-His compared to the human His-PRDX1 (Fig. 5C). The modifications with [D-Arg3], [D-Thr4], [D-Arg7], [D-Val8], [D-Ala9] RVA, and the cyclic 'native' RVA, did not improve binding to human PRX1-His (Fig. 5D-I). The negative control peptide showed no binding to human PRDX1-His (Fig. 5J). The mouse PRDX1-His showed very low millimolar affinity (KD) of 2560 ± 1340 µM to RVA (Fig. 5K) using SPR.

To further study the interaction between RVA peptide and PRDX1, we compared the binding of either mouse or human recombinant PRDX1-His proteins (PRDX1-His) to RVA or negative control peptide by using MST. For the MST study, the PRDX1-His proteins were labelled with fluorescence dye via the His-tag according to the manufacturer´s protocols. MST assays were performed using 16 premium capillaries filled with 10nM labelled target protein (PRDX1-His) in solution with serially diluted ligand (RVA or negative control peptide) ranging from 100 µM to 0.003 µM. The MST binding was measured at 40% excitation power, medium MST power and 25° C temperature using the Monolith NT.115 Pico instrument. Human PRDX1-His showed 3-fold higher nanomolar affinity to RVA (KD = 0.041 ± 0.016

μM) compared to its mouse ortholog (KD = 0.14 ± 0.10 μM) (Fig. 5L). Either human or mouse PRDX1-His showed no binding to the negative control peptide (Fig. 5M).

Results are shown in Table 5 and Table 6.

Table 5: SPR and MST binding affinities of human PRDX1-His to RVA or D-isomer derived variants.

RVA peptides	Binding Affinity, K_D (mean ± SD) μM	
Abbreviation	SPR	MST
RVA	86 ± 60	0.041 ± 0.016
[D-Cys^1] RVA	81 ± 65	
[D-Arg^2] RVA	223 ± 190	
[D-Thr^4] RVA	140 ± 110	
[D-Leu^5] RVA	49 ± 38	
[D-Arg^7] RVA	251 ± 142	
[D-Val^8] RVA	137 ± 94	
[D-Ala^9] RVA	312 ± 233	
Cyclic 'native' RVA	281	
Negative control	ND	ND

ND indicate no binding or not determined. K_D values are represented as mean ± SD.

Table 6: SPR and MST binding affinities of mouse PRDX1-His to RVA.

RVA peptides	Binding Affinity, K_D (mean ± SD) μM	
Abbreviation	SPR	MST
RVA	> 1000	0.14 ± 0.10
Negative control	ND	ND

ND indicate no binding or not determined. K_D values are represented as mean ± SD.

Example 8 - RVA-drug conjugate demonstrates cytotoxic effect *in vitro*

To study the RVA's ability to deliver drugs for targeted treatment of brain metastasis, we conjugated RVA to the antimicrobial peptide, $_D$(KLAKLAK)$_2$ (Ellerby et al. 1999) or to the antineoplastic agent, monomethyl auristatin E (MMAE) (Doronina et al. 2003) as a proof-of-concept. $_D$(KLAKLAK)$_2$ is a short peptide known to induce cell death after uptake into the cells by disrupting the mitochondrial membrane. However, $_D$(KLAKLAK)$_2$ is unable to internalize into cells by itself. Studies have documented the improved cytotoxic activity or efficacy of $_D$(KLAKLAK)$_2$ upon conjugation to tumour-targeting peptides.

Different tumour cells were treated with either RVA-conjugated or free $_D$(KLAKLAK)$_2$ at varying concentrations. After 24 hrs, cell viability was measured by using the MTT assay (Fig. 6). The IC_{50} value (concentration that kills 50% of the tumour cells at 24 hr) was determined from the dose response curves. The $_D$(KLAKLAK)$_2$-RVA demonstrated cytotoxic activity towards breast cancer cells as well as their brain metastatic counterparts. Brain

metastatic i231-BrM2 cells were the most sensitive (IC_{50} = 10.44μM, Fig. 6A), followed by 231 and 231-BrM2 (14.24μM and 15.03μM, respectively, Fig. 6C-D). The i231 cells showed the highest IC50 value (55.86μM, Fig. 6B) of these cell lines. The epithelial MCF10A cell line derived from human mammary gland showed 2-fold higher IC_{50} of 24.07μM (Fig. 6E) compared to the brain metastatic breast cancer cells. The $_D$(KLAKLAK)$_2$-RVA also killed very efficiently the melanoma derived brain metastatic cells (RALL.CB2, YDFR.CB3, and RKTJ.CB2) with IC_{50} values of 10.31μM, 18.97μM, and 12.67μM, respectively (Fig. 6F-H), as well as the patient-derived BT27 (IC_{50} = 17.83μM, Fig. 6I) and BT12 (IC_{50} =16.42μM, Fig. 6J) glioblastoma cells but showed no cytotoxicity towards the non-cancer cells (immortalized normal human astrocytes, NHA and human endothelial cells, HUART2, Fig. 6K-L). In addition, the human embryonic kidney cells, HEK293T, were unaffected by similar concentrations that killed half of the tumour cells (Fig. 6M). Table 7 shows the IC_{50} values of all the cell lines tested with $_D$(KLAKLAK)$_2$-RVA conjugates. The free $_D$(KLAKLAK)$_2$ did not affect the viability of any cell line tested. The cell death induced by the $_D$(KLAKLAK)$_2$-RVA conjugate was confirmed by the activation of cleaved-caspase 3 in the i231-BrM2 cells (Fig. 12).

Table 7: *In vitro* cytotoxic effect of the KLAK-RVA conjugate on cell viabilities of cancer and non-cancer cell lines.

Cell line	Cancer Type	Short name	IC_{50} (μM)
1	Invasive TN breast cancer-derived brain metastases	i231-BrM2	10.44
2	Invasive TN breast cancer	i231	55.86
3	TN breast cancer-derived brain metastases	231-BrM2	15.03
4	TN breast cancer	231	14.24
5	Human mammary epithelial cells	MCF10A	24.07
6	Human melanoma-derived brain metastases	RALL.CB2	10.31
7	Human melanoma-derived brain metastases	YDFR.CB3	18.97
8	Human melanoma-derived brain metastases	RKTJ.CB2	12.67
9	Patient-derived glioblastoma cells	BT27	17.83
10	Patient-derived glioblastoma cells	BT12	16.42
11	Normal human astrocytes	NHA	N.A.
12	Normal human endothelial cells	HUAR2T	N.A.
13	Human embryonic kidney cells	HEK293FT	N.A.

*N.A. indicate negligible or no cytotoxic activity.

To study further the specificity of the $_D$(KLAKLAK)$_2$-RVA conjugate, we conjugated the $_D$(KLAKLAK)$_2$ to the triple Ala (TrA) variant of the RVA that was designed based on the SAR data. In this peptide the three important amino acid residues (Cys^1, Leu^5, Gly^6) required both for the RVA internalization and its binding to the recombinant PRDX1 were replaced with Alanine residues to create a new control peptide AGRTAARVA (TrA negative control). The $_D$(KLAKLAK)$_2$-TrA control conjugate (blue curve in original image) showed significantly less cytotoxicity towards all tested cancer cells compared to the RVA-conjugate and failed to reach or reached barely the IC_{50} value with the concentrations used (Fig. 6A-D, 6F-H).

Example 9 - Clinical brain metastases samples show PRDX1 expression

We performed immunohistochemistry of 42 brain metastases (BM) from patients diagnosed with breast cancer (20), melanoma (20) and lung cancer (2) derived BMs, to study the clinical significance of PRDX1 (Fig. 7). All the 42 BM samples showed positive IHC staining for PRDX1 expression with varying staining intensities (low, intermediate, and high).

Example 10 - Pan-Cancer data analysis show higher PRDX1 expression in many tumour types

Pan-Cancer data comprising of 56,938 unique samples from GEO, GTex, TCGA and TARGET databases were analysed by TNM plot (Bartha and Győrffy 2021). PRDX1 showed significantly higher expressions in 16 different cancer indications out of 22 compared to normal tissues. Statistical analyses were performed by using Mann-Whitney U test and marked with red asterisks (Fig. 8).

Example 11 - Peptide-targeted drug (monomethyl auristatin E or F =MMAE or MMAF, respectively) efficiently kills triple negative breast cancer (TNBC) -derived brain metastatic cells in vitro.

The TNBC-derived brain metastatic cells (i231-BrM2) (Fig 13A panel A, Fig 13B panel C) or primary TNBC cells (i231) (Fig 13A panel B, fig 13B panel D) were treated with different concentrations of RVA- or control peptide (TrA neg. ctrl) -conjugated drug molecules. Cell viability was measured after 24 hrs incubation and the IC_{50} values were determined. Targeted RVA-conjugated drug conjugates showed higher potency compared to the non-targeted control peptide conjugated drug conjugates. Even though both primary and brain metastatic cells were killed by the targeted drug conjugates, the brain metastatic cells were killed more efficiently than the primary TNBC cells (40-fold difference with MMAE and 3,9-fold difference with MMAF).

Example 12 - Peptide-targeted drug (toxic antimicrobial peptide, $_D$(KLAKLAK)$_2$) efficiently inhibited brain metastatic tumor progression in vivo.

Immunocompromised mice were implanted intracranially with TNBC-derived brain metastatic cells. At day 7 post-implantation the mice were randomized to four different treatment groups (n = 10 in each group) and the treatment was started (Fig. 14A). Mice were treated every other day by intravenous injections of vehicle (Fig. 14C), control peptide conjugated $_{D}(KLAKLAK)_{2,}$ thereafter referred to as KLAK (KLAK-TrA, 5 mg/kg) (Fig. 14D), free KLAK (5mg/kg) (Fig. 14E) or targeted peptide-conjugated drug (KLAK-RVA, 5 mg/kg) (Fig. 14F). Tumor growth was monitored by whole-body bioluminescence imaging to visualize the luciferase-expressing cancer cells (Fig. 14B). The RVA-targeted drug significantly inhibited the tumor growth compared to any control group (Fig. 14B, Fig. 14C - Fig. 14F).

Example 13 - Peptide-targeted drug (toxic antimicrobial peptide, KLAK) efficiently inhibited brain metastatic tumor progression in vivo.

After the treatment, mice were euthanized, and brains were excised and prepared for histological analysis (Figure 15, panels A-D). Representative micrographs of the mice brains stained with antibodies against human vimentin (red in the two left most panels), anti-PRDX1 (receptor for the RVA-peptide, green in the second to right most panels, anti-CD31 (blood vessel marker, red green in the second to most panels) or apoptotic, dead/dying cells (TUNEL, red in the right most panels). All panels following the left most panels, show higher magnification of the boxed area in the left most panel. Quantification of the total tumor area in the mouse brains (Fig. 14E). Quantification of the percentage of dead/dying cells (TUNEL+) (Fig. 15F). The histological analyses confirm the results of the bioluminescence imaging and show significantly reduced tumor size and increase in dying tumor cells in the RVA-peptide targeted mice compared to the control groups.

Example 14 - Peptide-targeted [^{18}F]RVA clearly visualizes the brain metastatic tumors in preclinical mice model of human brain metastasis. Figure 16.

Panel A. Representative PET-CT image of the whole mouse bearing metastatic brain tumors. Peptide-derived radioactive signal is seen in the tumor (red box) and in kidneys (white box) via which the peptide excreted from the body. Panel B. Shows the peptide-mediated PET-imaging of metastatic brain tumors after intravenous injection (white arrow points to the tumor), while no accumulation of the radioactivity is detected in the healthy brain (Panel C). Panel D. The standardized uptake value (SUV) shows the higher uptake and retention of the radioactive RVA-peptide in the brain metastasis compared to the healthy brain following intravenous injection of the peptide.

Example 15 - NAVIGATOR for PET imaging of bone and lung metastases in intracardiac mouse model. Figure 17.

Peptide-targeted [^{18}F]RVA clearly visualizes lung and bone metastasis in addition to the brain metastasis in preclinical mouse model. Fig. 17 A – Fig. 17B.

The whole-body bioluminescence (BLI) and the PET/CT images of multiorgan metastasis-bearing (Fig. A) or healthy (Fig. B) mouse. BLI imaging visualizes tumors cells while the PET-imaging shows the uptake of the radioactive peptide. White arrows mark the metastatic sites containing radioactive peptide in A and corresponding sites in healthy mouse without signal from intravenously injected radiolabelled peptide. Fig. C. Higher magnification images of the boxed areas in A. Arrows mark the deteriorated bone due to metastatic growth in the right leg (R), which is not visible in the left leg (L) of the same animal that does not contain any tumor cells. Fig. 17D. Higher magnification images of the lung metastasis showing the uptake of radioactive peptide.

Example 16 - Pancreatic tumour tissue (PDAC xenografts) autoradiography. [^{18}F]RVA binds strongly to human pancreatic adenocarcinoma xenografts.

Human pancreatic adenocarcinoma organoids were implanted in the immunocompromised mice via the splenic vein to create orthotopic tumors in the pancreas. Sections were prepared from the xenografts and the radiolabeled [^{18}F]RVA peptide was incubated on the sections. After extensive washes the autoradiography revealed strong binding in all tumor sections derived from three different mice. Results are shown in Fig. 18.

Example 17 - Clinical Breast Cancer brain metastases autoradiography. [^{18}F]RVA binds very specifically.

A. Sections of human clinical breast cancer brain metastasis bound strongly the [^{18}F]RVA peptide. B-C. When cold RVA-peptide (1 or 10 μM) was added to the sections prior the addition of the [^{18}F]RVA peptide, binding of the radioactive peptide significantly reduced. This indicates that the radiolabelled peptide binds very specifically to its target protein. Results are shown in Fig. 19.

Example 18 - Clinical Breast Cancer brain metastases autoradiography. [^{18}F]RVA peptide binds specifically to human clinical breast cancer brain metastasis in the areas with high target protein expression. Figure 20.

A. Sections of human clinical breast cancer brain metastasis bound strongly the [^{18}F]RVA peptide. B. When the consecutive sections were stained with antibodies against PRDX1 (RVA target protein), high PRDX1 (green) expression localized to the areas that showed high radiolabel uptake. C. Dead/dying cell were visualized by the TUNEL staining (red). Majority of the tumor cells are alive and thus, able to uptake the radiolabelled peptide.

CLAIMS

1. A homing peptide, or its functional variant, comprising an amino acid sequence CxxxLG.
2. The homing peptide or its functional variant according to claim 1 comprising an amino acid sequence CxxxLGxxx.
3. The homing peptide or its functional variant according to claim 1 comprising an amino acid sequence CGRxLG, SEQ ID NO: 53; or CGRxLGxxx, SEQ ID NO: 54; or CGRxLGRVA, SEQ ID NO: 55.
4. The homing peptide or its functional variant according to claim 1 comprising an amino acid sequence CxRxLGRVA, SEQ ID NO: 3.
5. The homing peptide or its functional variant according to any one of claims 1-4 comprising an amino acid sequence CGRTLGRVA, SEQ ID NO: 4.
6. The homing peptide or its functional variant according to any one of claims 1-5, wherein the homing peptide has no more than 30 amino acids, preferably 6-15 amino acids.
7. The homing peptide or its functional variant according to any one of claims 1-6 having C as the N-terminal residue and optionally A as the C-terminal residue.
8. The homing peptide or its functional variant according to any one of the claims 1-7 capable of specifically binging PRDX1 (peroxiredoxin 1), preferably an N-terminal domain of PRDX1.
9. A conjugate comprising the homing peptide or its functional variant of any one of the claims 1-8 and a cargo molecule, wherein the cargo molecule is at least one of: a nucleic acid, antisense molecule, aptamer, ribozyme, triplex forming oligonucleotide, external guide sequence, RNAi, CRISPR/Cas, zinc finger nuclease, transcription activator-like effector nucleases, an expression vector encoding a protein or a functional nucleic acid, a vectors suitable for integration into a cells genome or expressed extra-chromosomally, antibody, monoclonal antibody, tagged molecule, fluorescence tag molecule, label, nanoparticle, conjugated nanoparticle, liposomal payload, small molecule, anti-cancer agent, an anti-inflammatory agent, an immunomodulatory agent, an antimicrobial agent, an organic compound, an organometallic compound; a hydrophilic compound, a hydrophobic compound, an amphiphilic compounds, toxic agent, radioisotope, a radionuclide, a fluorescent tag/compound/dye, a magnetic tag/molecule, paramagnetic molecule, x-ray imaging agent, contrast media, dye, near infra-red dyes, SPECT imaging agent, PET imaging agent, peptide, lipid, glycolipid, glycoprotein, polypeptide, a therapeutic agent, a diagnostic agent, a prophylactic agent, a nutraceutical agent.

10. The conjugate of claim 9, wherein the cargo molecule is an imaging moiety or a diagnostic moiety, alone or in combination with one or more therapeutic moiety.

11. The conjugate of claim 9, wherein cargo molecule is an imaging agent.

12. A pharmaceutical composition comprising the conjugate of any one of the claims 9-11 and a pharmaceutically acceptable carrier.

13. The homing peptide or its functional variant according to any one of claims 1-8, or the conjugate of any one of claims 9-11, for use as a medicine.

14. The homing peptide or its functional variant according to any one of claims 1-8, or the conjugate of any one of claims 9-11, for use in treating cancer, preferably a brain metastasis.

15. A method of *in vivo*, or *ex vivo*, or *in vitro* diagnosis of cancer using the homing peptide or its functional variant according to any one of claims 1-8, or the conjugate of any one of claims 9-11, wherein the cancer is preferably at least one of melanoma, breast carcinoma, lung carcinoma, brain cancer, brain metastatic cell, or any other cancer with PRDX1 expression.

16. A method of delivering a cargo molecule into a cancer cell comprising contacting the cancer cell with the conjugate of any one of claims 9-11.

17. A method of identifying a cancerous target cell or cancerous target tissue in a biological sample, the method comprising:

(a) contacting the biological sample with the conjugate of any one of claims 9-11 for a time and under conditions sufficient for specifically binding the conjugate to the target cell or the target tissue; and

(b) detecting the presence or absence of the bound conjugate in the biological sample;

wherein presence of the conjugate in the detecting step (b) indicates that the cancerous target cell or the cancerous target tissue is present in the biological sample.

18. The method of claim 17, wherein the step (b) is carried out by immunohistochemistry, fluorescent imaging, radioimmunoassay, or immunofluorescence.

19. A kit comprising: (a) a container comprising a pharmaceutical composition containing the homing peptide or its functional variant according to any one of claims 1-8, the conjugate of any one of claims 9-11, or the pharmaceutical composition of claim 12, in solution or in a lyophilized form; (b) optionally, a second container containing a diluent

or reconstituting solution for the lyophilized formulation; and (c) optionally, instructions for use of the solution or reconstitution and/or use of the lyophilized formulation.

20. A complex comprising the homing peptide or its functional variant according to any one of the claims 1-8 and a target molecule.

5