Molecular ZIP codes in targeted drug delivery

Erkki Ruoslahti

Edited by Richard Hynes, Massachusetts Institute of Technology, Cambridge, MA; received January 27, 2022; accepted May 5, 2022

The term “molecular ZIP (or area) codes” refers to an originally hypothetical systems of cell adhesion molecules that would control cell trafficking in the body. Subsequent discovery of the integrins, cadherins, and other cell adhesion molecules confirmed this hypothesis. The recognition system encompassing integrins and their ligands came particularly close to fulfilling the original ZIP code hypothesis, as multiple integrins with closely related specificities mediate cell adhesion by binding to an RGD or related sequence in various extracellular matrix proteins. Diseased tissues have their own molecular addresses that, although not necessarily involved in cell trafficking, can be made use of in targeted drug delivery. This article discusses the molecular basis of ZIP codes and the extensive effort under way to harness them for drug delivery purposes.

What are “molecular ZIP codes”? The origins of this concept go back to the search for molecular explanation for precision and complexity of the formation of body structure during embryonal development. Observations such as spontaneous sorting of like from unlike sponge cells (1) and the navigation by growing retinal axons to their predetermined connections in the brain tectum (2) led to a hypothesis that “area code” molecules at cell surfaces guided the cells of the embryo to the appropriate place in the developing organism (3). Subsequent discovery of molecules that mediated cell–extracellular matrix and cell–cell recognition, provided a molecular basis for the hypothesis (4–6). The definition of molecular ZIP codes (ZCs) has since expanded beyond developmental biology and currently encompasses any molecule that the expression of which in the adult body is restricted to a certain cell type, tissue, organ, or pathological condition. Also, the original metaphor of area codes in the telephone system has become ZCs of the postal system. As the molecular ZIP (or area) code, can be related to the ZIP code. The vascular endothelium displays a remarkable degree of tissue- and disease-specific molecular diversity. It provides molecular addresses for the trafficking of cells such as lymphocytes (19) and metastatic tumor cells (20, 21) (Fig. 1).

Analyses of vascular heterogeneity in situ by in vivo screening of phage libraries has yielded peptides that selectively recognize the blood vessels of a single organ (Fig. 1), suggesting that many, if not all, tissues put a signature on their vasculature that make that vascular bed unique in the body (24). So far, one vascular receptor for such organ-specific peptides has been identified molecularly. It is a membrane dipeptidase involved in the metabolism of glutathione, leukotriene D₄, and certain lactam antibiotics (25). This finding, and the fact that many kinds of molecules are represented among disease-related vascular markers

ZC Molecules

Among the ZC molecules (ZCMs), cell adhesion proteins mostly closely confirm to the ZC concept. Integrins are a large family of genetically related proteins that recognize the RGD and related motifs in their protein ligands, such as the extracellular matrix proteins fibronectin, vitronectin, collagens and laminins (4, 5, 11, 18). Cadherins, Eph receptors and ephrins, and various immunoglobulin-G–domain cell–cell adhesion proteins similarly mediate cell–cell adhesion (6). However, I am using the ZC term broadly to include any molecule that is selectively expressed in a tissue or disease, regardless of whether it is known to have a function in development or cell trafficking.

The vascular endothelium displays a remarkable degree of tissue- and disease-specific molecular diversity. It provides molecular addresses for the trafficking of cells such as lymphocytes (19) and metastatic tumor cells (20, 21) (Fig. 1).

Analyzes of vascular heterogeneity in situ by in vivo screening of phage libraries has yielded peptides that selectively recognize the blood vessels of a single organ (Fig. 1), suggesting that many, if not all, tissues put a signature on their vasculature that make that vascular bed unique in the body (24). So far, one vascular receptor for such organ-specific peptides has been identified molecularly. It is a membrane dipeptidase involved in the metabolism of glutathione, leukotriene D₄, and certain lactam antibiotics (25). This finding, and the fact that many kinds of molecules are represented among disease-related vascular markers

Author contributions: E.R. wrote the paper.

The author is a founder, shareholder, and board of directors chairman of Cend Therapeutics, a clinical stage company developing targeted cancer treatments, and of Alvoco, a company targeting brain diseases.

This article is a PNAS Direct Submission.

Copyright © 2022 the Author(s). Published by PNAS. This article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

1To whom correspondence may be addressed. Email: ruoslahti@sbpdiscovery.org. Published June 30, 2022.
Some of the ZCs may be brought to growing vessels by bone marrow-derived cells that contribute to the newly formed endothelium (32). This appears to be the case with the expression of p32/gC1qR on the surface of endothelial cells in tumor blood vessels and lymphatics (33). p32/C1qR is a marker of activated macrophages and myeloid cells, and bone marrow-derived myeloid cells may be the source of p32 on tumor vessels (33).

Despite the sharing of angiogenesis ZCs by many diseases, differential targeting of individual diseases involving angiogenesis may be possible, at least to a degree. There are quantitative differences in the expression levels of individual markers between physiological and pathological (tumor) angiogenesis as measured by gene expression analyses (34). Moreover, phage library screening targeting skin and tendon wounds yielded a different complement of angiogenesis-targeting peptides than screening with tumors (35).

The angiogenesis ZCs that are most frequently used in targeting drugs to tumors are integrins αvβ3 and αvβ5. These integrins, along with α5β1, are abundantly expressed in tumor vessels and essentially absent in normal vessels (36, 37). They recognize the three-amino-acid RGD motif in their ligand proteins, and short RGD-containing peptides (11) can be used to direct drug to tumors (38). The effective targets of the RGD peptides are the two αv integrins; the affinity of α5β1 for the peptides is too low for this integrin to contribute to the targeting. There is a great deal of literature from the past 25 y on the use of RGD-containing peptides in integrin-directed targeting (39–41). Even the iRGD peptide, the prototype compound among the more recently discovered tumor-penetrating peptides, homes to tumors using its RGD motif (42); see below). Other vascular ZCMs used in tumor targeting include prostate-specific membrane antigen (PSMA) (43), vascular endothelial adhesion molecule-1 (V-CAM-1) (44), nucleolin (45, 46), a form of aminopeptidase N (47, 48), ephrin-A1 ligand and its receptor EphA2 (49), and p32/gC1qR (33). In the case of p32/gC1qR, both the blood vessel and lymphatic endothelia are positive for cell surface expression of this protein. Like the tumor-associated integrins, these proteins are expressed in the vasculature of most cancers, and this is also true of PSMA, despite its name.

The mechanisms that make vascular ZCMs tumor specific vary. The most obvious mechanism, overexpression at mRNA and/or protein levels, accounts for only some of the selectivity. The angiogenesis-associated integrins αvβ3, αvβ5, and α5β1 belong to this category. However, nucleolin and p32/gC1qR become selective for tumor endothelium (and other cell types in tumors) because they are confined to the cell interior in normal cells but are also present at the cell surface in tumors (33, 45, 50). The surface presence makes them available for the binding of targeting probes in tumor but not in normal tissues. Overexpression in tumors also contributes to the selectivity, at least in the case of p32/gC1qR (33, 51). Nucleolin and p32/gC1qR appear to be cycling between the interior and surface of cells, and this may also be the case with the nuclear receptor RXRB, which is a marker of tumor-associated macrophages (52). Thus, the aberrant surface expression of intracellular molecules appears to be a common phenomenon in tumor-associated and other activated cells. The reasons for this phenomenon are not understood.
but its selectivity and payload-internalization potential provide a useful tool for targeting drugs to tumors (38).

**ZCMs in Tumor Cells and Tumor Stromal Cells**

ZCMs are often shared by more than one cell type in tumors. A case in point is the αv integrins, which are expressed in tumors by endothelial cells (36), tumor fibroblasts, and tumor cells (53). The p32/gC1qR protein is expressed at the cell surface in tumor endothelial cells, tumor macrophages, and tumor cells (32). The expression in tumor macrophages is particularly high, which makes targeting of this protein largely specific for these cells. PSMA is expressed by tumor cells and tumor endothelial cells in prostate cancer but is only present in the tumor endothelial cells of most other types of tumors (43). As its name indicates, the fibroblast-activated protein is expressed by fibroblasts that have been activated by various biological processes, such as inflammation or cancer. A peptide that binds to this protein has been used for drug delivery to cancer-associated fibroblasts (54).

**ZCMs in the ECM**

The ECM in tumors is altered (55, 56), and some of these changes can serve as ZCs for tumor targeting. A ZC can arise from an ECM protein in two ways: by an increase in an ECM protein or a specific splice variant, or by increased accessibility to a probe of a protein or protein epitope in tumor ECM. Fibronectin and tenascin are among the ECM proteins with multiple splice variants (57, 58). Some of these are predominantly expressed in embryonic and fetal tissues, and their re-expression in tumors makes them targets for drug delivery into tumors (58–61).

A peptide originally identified by a combination of screening on Matrigel in vitro and tumors in vivo appears to bind to laminin-nidogen complexes in tumor ECM (63). These complexes are highly expressed in tumors, but increased exposure of the epitope that binds the peptide may also play a role in the tumor specificity. An ECM-binding peptide that homes to brain injuries also derives its specificity from both elevated expression and increased exposure of the target protein (63). The reason for the increased availability of the target for the peptide in this case is that brain injury causes disruption of the blood-brain barrier, making the ECM in the injured brain accessible to the peptide. ECM-bound connective tissue growth factor is a marker of blood vessels and activated astrocytes in Alzheimer’s disease and other degenerative brain diseases (64).

Fibrin is a unique member of the target molecules in the tumor ECM because there is subtle blood clotting in tumor vessels (65). The increased clotting is a result of tissue factor expression on the luminal surface of tumor endothelial cells, which is absent in normal endothelium (66). Moreover, the leakiness of tumor vessels also allows some fibrinogen to enter the extravascular tumor matrix, where it encounters tissue factor and is converted into fibrin. The fibrin deposits in tumor vessels and tumor ECM provide a specific target for peptide and antibody-mediated drug delivery (67–70). A bioinspired amplification system has been designed in which subtly thrombogenic, peptide-guided nanoparticles bind to the fibrin deposits in tumor vessels and amplify their own binding by causing more clotting, which eventually occludes the vessels (72). The system is effective in curtailing tumor growth and could likely be made even more effective by adding an anticancer drug to the nanoparticles. The fear of systemic clotting has prevented further development of the system. Among the ECM targets, the extra domain A fibronectin splice variant has reached advanced clinical trials as a targeted cancer therapy (72).

The major advantage of ECM targeting is that ECM is an abundant component of tumors, particularly the difficult-to-treat fibrous ones, and a ligand that binds to tumor ECM allows a larger amount of a drug to be specifically delivered than the less abundant cell-surface molecules (see Quantitative Aspects of Targeted Drug Delivery). Targeting of drugs to tumor ECM would appear to have the disadvantage that the payload is brought to the matrix and not to the cells that are the actual target of the drug. However, ECM binding may not be a problem because of two factors: Once the drug has become concentrated in the tumor through the matrix binding, it can diffuse to the cells. Moreover, ECM components are often internalized by cells and can take a bound peptide and its payload with them (56, 58, 63, 64).

**Tumor-Penetrating Peptides and Drug Delivery through the CendR Pathway**

Certain tumor-homing peptides penetrate deep into tumors by using an endocytic transcytosis pathway that transports the peptide across the wall of tumor vessels through tumor stroma into tumor cells. The transport mechanism has been elucidated and consists of three steps: In the first, the tumor-penetrating peptide initially binds to a receptor expressed specifically on the luminal surface of tumor endothelium through a sequence motif such as RGD. In the second step, a proteolytic cleavage at the endothelial surface causes another sequence motif, RXXR/K, to become the C terminus of the peptide, which, when in the C-terminal position, enables binding to a new receptor, NRP-1. As the NRP binding requires that the RXR/K sequence motif is at the C terminus of the peptide, it has been named the CendR motif, and the transport pathway activated by the NRP binding is referred to as the CendR pathway (10) (Fig. 2).

The prototype tumor-penetrating peptide iRGD (42) contains the RGD motif and initially binds to αv integrins on tumor endothelium. Other peptides can use a different first receptor but like iRGD, they are converted into an NRP-binding CendR peptide. The initial receptor for the LyP-1 peptide (75) is p32/gC1qR on tumor endothelium (33). New tumor-penetrating peptides can be created de novo by following the same design. Peptides containing the NGR sequence motif home to tumors by recognizing aminopeptidase N, perhaps some kind of a tumor-specific variant of this enzyme, and home to tumor vessels (47, 48). An NGR peptide was converted to a tumor-penetrating peptide by adding another arginine to create the CendR motif RNGR (76, 77).

The protease (or proteases) that exposes the CendR motif in tumor-penetrating peptides is likely to be a furin or furin-like enzyme because these proprotein convertases cleave their target proteins after a CendR motif (7). However, any protease that cleaves after a basic residue (arginine or lysine)
can convert a peptide or protein containing a CendR motif into an active NRP binder. Examples include urokinase and trypsin (78). The NRP binding of the RXXR/K triggers the activation of an endocytic transcytosis and transtissue transport pathway that distributes the peptide throughout tumor tissue (10, 42, 79). Both NRP-1 and NRP-2 can act as the receptor that activates this pathway (80). Loosening of endothelial junctions has been reported to be another consequence of NRP stimulation and was made use of in promoting immune cell delivery to tumors (81, 82).

The specificity of peptides such as iRGD, LyP-1, and iNGR arises primarily from binding to their primary receptor at the vascular endothelium. Fig. 3 shows and example of the tumor-homing specificity of the tumor-homing CendR peptide LyP-1. The αv integrins that serve as the primary receptors for iRGD are specifically expressed in tumor vessels. They are not detectably expressed in normal endothelium but are expressed in a variety of nonmalignant conditions that involve angiogenesis, such as tissue regeneration and inflammation (84, 85). The LyP-1 and iNGR receptors p32/gC1qR and aminopeptidase N, respectively, are similarly expressed in tumor endothelial cells and not in normal endothelium. Increased tumor expression of furins, the proteases that are the likely processors of the tumor-penetrating CendR-motif peptides (86), may add to the tumor specificity. NRP-1 is ubiquitously expressed in normal tissues, including endothelial cells, but it is overexpressed in tumors, such that aggressive tumors tend to express the highest levels of it. The tissue distribution of NRP-2 is more restricted but also present in higher levels in tumors (87, 88). Thus, the NRPs may also contribute to the tumor specificity of the tumor-penetrating peptides. Peptides that bind directly to NRP have also been used in delivering drugs to tumors (89–91), but they lack the specificity of the peptides that target receptors specific for tumor vessels and are then converted there to NRP binders. An exception may be a short peptide called tLyP-1, which has a C-terminal CendR motif that binds to NRPs, but nonetheless seems to be specific for tumors (80). It is possible that this peptide also binds to a second, as yet unknown receptor.

A remarkable property of the peptides with the tumor-penetrating CendR motif is that they can deliver into a tumor a payload that is coadministered with the peptide, not conjugated to it (22, 92). The CendR endocytosis resembles macropinocytosis. The main differences in macropinocytosis are that the activation of the CendR pathway is receptor dependent, whereas classical macropinocytosis is constitutive, and that CendR transport is not affected by pharmacological macropinocytosis inhibitors. Also, CendR is a transcytosis process (22, 93). The CendR endocytic vesicles are approximately 200 nm in diameter, and as they form at the cell surface, they take up fluid surrounding the cell, including any drug present in it. The large size of the vesicles allows them to transport payload as large as nanoparticles.

The coadministration mode of using tumor-penetrating peptides provides significant advantages: Once a peptide is shown to be safe, any approved drug can be coadministered with it, whereas each drug-peptide conjugate is a new chemical entity that requires its own validation. Thus, the coadministration mode greatly simplifies the manufacturing and regulatory path into the clinic. Moreover, the delivery capacity of the system is not limited by the quantity and accessibility of the receptor used in a coupled delivery system, which often is a limiting factor (see Quantitative Aspects of Targeted Drug Delivery).

**ZCMs in Nonmalignant Diseases**

Many nonmalignant diseases are also known to possess ZCs. Examples include V-CAM expression in the endothelium of vessels in tissues activated by processes such as inflammation. The angiogenesis-related makers discussed above in the context of tumor vessels are generally also up-regulated in nonmalignant conditions that involve the
formation of new blood vessels. A prominent example of pathological angiogenesis in a nonmalignant disease is macular degeneration, in which RGD peptides have been proposed as a potential treatment (94). Other disease-related ZCs include a tenasin-related ECM epitope in brain injuries (65) (Fig. 3) and a change in the brain vessels of Alzheimer’s disease, the expression of connective tissue growth factor (66). A heparan sulfate epitope related to angiogenesis and tissue regeneration has been used to target muscular dystrophy lesions (95).

Quantitative Aspects of Targeted Drug Delivery

An important aspect of ligand-directed drug targeting that has not received enough attention relates to the quantity and availability of the targeted ZCMs in the relevant tissue. This issue is important, because the number of receptor molecules available limits the amount of drug that can be specifically delivered. If the amount of a drug conjugate is greater than the available receptors can accommodate, then the excess remains untargeted. If the excess is substantial, then the effect of specific targeting will be drowned out by the untargeted background.

Direct quantification of two tumor receptors, αv integrins and Her2, showed between 100 and 1800 pmol/g tumor of these receptors in tumors generated with three commonly used cell lines (96). Only a fraction (4–6%) of these receptors was available for the quantification probe. The availability of the target molecule for the probe could be increased by about four- to fivefold by coadministering the probe with the tumor-penetrating IRGD peptide. As the time the tissues of the tumor mouse were exposed to the circulating probe in these experiments was short, these numbers likely underestimate the number of receptors. Even so, they are soberingly low. At a 1:1 ligand-to-receptor ratio, the amount of drug that could be targeted would be well below 1 nmol/g tumor, which would be insufficient for most anticancer drugs. A peptide or other targeting ligand is likely to have nanomolar to low micromolar affinity for its receptor, making it necessary to administer more of the targeted conjugate to achieve receptor saturation than would be needed under ideal binding conditions. There are several potential solutions to this problem. One is using high-affinity ligands, such as antibodies or aptamers (97), conjugated to a drug with high specific activity that reduces the amount of conjugate needed. A potential downside with high-affinity ligands is the so-called binding site barrier phenomenon (98, 99), which reduces the penetration of high-affinity ligands into tumors. Self-amplifying targeting systems (73, 100, 101) provide another possible solution but have not advanced beyond proof of concept of the approach.

Nanoparticles provide an ideal vehicle for ligand-targeted delivery (40), particularly for nucleic acid drugs, which often require protection from degradation (102, 103). The presence of multiple ligand molecules on a single nanoparticle enables multivalent binding to target receptors, circumventing problems with low affinity, particularly of peptides. Also, more drug per target receptor can be loaded in or on a nanoparticle than in 1:1 conjugates (38). Finally, coadministering a drug with a tumor-penetrating peptide that activates the CendR pathway bypasses the need to use conjugates. Tumor receptors are only used to activate the pathway, and the rate-limiting factor becomes the capacity of the CendR pathway, which is likely to be higher than the capacity of the equivalent conjugate approach.

Ligand-Directed Drug Delivery in the Clinic

Antibody-drug conjugates directed to target molecules on cancer cells are already in clinical use for some solid tumors (16). A significant challenge in this field is poor tissue penetration, which limits the delivery of drugs into tumor tissue. In solid tumors, the tumor cells are embedded in connective tissue, which in desmoplastic tumors is extremely dense and forms a formidable barrier to drug entry. Another obstacle is the high hydrostatic pressure of tumor tissue that pushes tissue fluid, and any drug in it, out of the tumor (104, 105). Antibody-drug conjugates are relatively large, which exacerbates the penetration problem. Nanoparticle delivery vehicles, which are intensively studied in experimental cancer treatments, face the same problem. The smaller size of peptide–drug conjugates helps them penetrate tissue, but the relatively low affinity of monovalent peptide–receptor interactions tends to limit
the effectiveness of such conjugates. Nevertheless, a multitude of antibody and peptide conjugates and various kinds of drug-carrying nanoparticles are at various stages of clinical trials. A potential solution to the penetration problem is breaking down the principal barrier, the tumor ECM. Clinical trials on this approach have been disappointing (106, 107). The tumor-penetrating peptide technology provides another possible solution for the tumor-penetration problem. These peptides activate the CendR transcytosis transport pathway in a tumor-specific manner, which converts tumor ECM from an obstacle to a conduit for drug delivery (53). The prototype cell penetrating peptide, iRGD, has been through a phase 1/2a clinical trial as an enhancer in pancreatic cancer chemotherapy [NCCT03517176 (108)] and has entered phase 2b trials (https://clinicaltrials.gov/ct2/results?cond=&term=CEND-1&cntry=&state=&city=&dist=).

Conclusions

Disease-specific ZCMs represent attractive targets for ligand-directed drug delivery, especially in cancer. The vascular endothelium of tumor vessels, the tumor cells themselves, and tumor stromal cells all express ZCMs with various degrees of tumor specificity that can be used to increase the selectivity of anticancer treatments. The ECM of tumors also carries ZC epitopes. Other disease conditions, such as wounds, brain injuries, Alzheimer’s disease, and atherosclerotic plaques can be similarly targeted. The technology has been successfully brought to the clinic in the treatment of some cancers. The main obstacle of drug delivery to solid tumors and other diseases requiring delivery into tissue is tissue penetration of drugs, particularly in tumors, as well as nanoparticles and antibodies. Ongoing efforts to address the tumor penetration issue, if successful, could greatly increase the utility of active drug targeting for cancer treatment, with new treatments for other diseases to follow.

Data Availability. There are no data underlying this work.

ACKNOWLEDGMENTS. I thank Dr. Eva Engvall for comments on the manuscript.
et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.