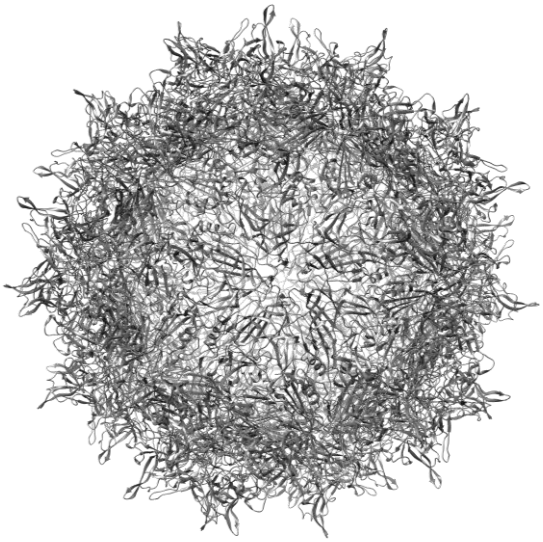
Design and characterization of EGFR-specific peptides and re-targeted rAAVs for tumor therapy



Rebecca Christine Feiner

Bielefeld

- 2019 -

Design and characterization of EGFR-specific peptides and re-targeted rAAVs for tumor therapy

Dissertation

Zur Erlangung des akademischen Grades

Doctor rerum naturalium (Dr. rer. nat.)

Zelluläre und Molekulare Biotechnologie

Technische Fakultät

Universität Bielefeld

vorgelegt von

Rebecca Christine Feiner

Bielefeld

2019

Die vorliegende Arbeit entstand in der Zeit von

Februar 2016 bis Mai 2019

in der Arbeitsgruppe

- Zelluläre und Molekulare Biotechnologie -

an der Technische Fakultät der Universität Bielefeld

unter Leitung von

Herrn Prof. Dr. Kristian M. Müller

1. Gutachter: Prof. Dr. Kristian M. Müller
   Zelluläre und Molekulare Biotechnologie, Technische Fakultät
   Universität Bielefeld

2. Gutachter: Prof. Dr. Norbert Sewald
   Organische und Bioorganische Chemie, Fakultät für Chemie
   Universität Bielefeld

3. Gutachter: Prof. Dr. Stefan Kochanek
   Abteilung für Gentherapie,
   Universitätsklinikum Ulm

Danksagung

An erster Stelle möchte ich mich ganz herzlich bei meinem Doktorvater Prof. Dr. Kristian M. Müller für die Aufnahme in seine Arbeitsgruppe bedanken, die mir die Möglichkeit gab an diesem spannenden, herausfordernden Thema arbeiten zu können. Vielen Dank für den großen Freiraum, den du mir bei der Gestaltung meines Projekts gewährt hast und die zahlreichen wertvollen Ratschläge und Diskussionen, die maßgeblich zum Erfolg der Arbeit geführt haben.

Prof. Dr. Norbert Sewald danke ich sehr für die Übernahme des Zweitgutachtens dieser Arbeit und die Unterstützung während des Studiums in Master- sowie Bachelorarbeit.

Prof. Dr. Stefan Kochanek au Ulm danke ich sehr für die Übernahme des externen Gutachtens.

Dem gesamten Arbeitskreis Zelluläre und Molekulare Biotechnologie danke ich für eine einmalige Zeit in einem tollen Arbeitsumfeld. Zahlreiche AG Ausflüge und Weihnachtsmarktbesuche als auch unzählige, teils selbstgebraute, Feierabendbiere werden mir in guter Erinnerung bleiben. Ein besonderer Dank geht dabei an Philipp Borchert, der mich bei unterschiedlichsten Experimenten unterstützt hat und immer mit Rat und Tat zur Seite stand. Auch ein großer Dank geht an Ulrike Baillie für die großartige Unterstützung neben dem üblichen Laborwahnsinn.

All denjenigen, die im Rahmen ihrer Abschlussarbeiten, Ausbildung und Forschungsprojekte für mich tätig waren, danke ich für ihr großes Engagement und die Mithilfe zu den schönen Ergebnissen, die in dieser Arbeit entstanden sind. Danke, Lina Schütte, Daniel Gödde, Lennard Karsten, Claire Rothschild, Julia Niehues, Irina Schierbaum, Greta Bischof, Ina Pennè, Thilo Pohle und Christopher Lindenkamp.

Bei Dr. Lea Krutzke möchte ich mich für die gute Zusammenarbeit im Rahmen der *in ovo* Experimente bedanken.

Meinen Labor- und Bürokollegen Kathrin, Julian, To, Marco und Georg möchte ich hiermit noch einmal besonders danken. Durch die vielfältigen, hilfreichen Diskussionen und die tolle Arbeitsatmosphäre ist die Zeit im Labor und im Büro immer wie im Flug vergangen.

Dem Chemikalienlager der Fakultät für Chemie, im Besonderen Heike Koselleck, danke ich für die zahlreichen Angebote, die herausgesucht wurden, um uns mit möglichst preiswerten Produkten zu versorgen. Weiterhin danke ich Heinrich aus der mechanischen Werkstatt und Lothar für die Reparatur manch aussichtlos-kaputter Elektrogeräte.

Großer Dank an Hendrik und Eduard aus der OCIII für die Unterbringung in ihren Laboren während der Peptidsynthese, sowie zahlreiche Messungen an LC/MS und MALDI. Ein besonderer Dank geht darüber hinaus an Isabell, die mich durch die Synthese der Peptide, der zweiten Generation, außerordentlich unterstützt hat.

Viel Unterstützung erhielt ich während langer Mikroskopie-Tage auch von Thorsten Seidel am LSM Mikroskop. An dieser Stelle möchte ich auch den Biochemiearbeitsgruppen danken, dass ich neben der liebgewonnenen Ultrazentrifuge auch die French Press, den Tecan Reader und diverse andere Geräte nutzen durfte.

Weiterhin geht ein großes Dankeschön an die AG Zellkulturtechnik, die mir nicht nur die Bioreaktorwelt nähergebracht hat, sondern auch durch zahlreichen interessanten Gespräche auf dem E2-Flur meine Arbeit unterstützt hat.

Ich möchte mich weiterhin bei meinen Kommilitonen und natürlich auch denjenigen, die mich abseits des Labors unterstützt haben bedanken. Hier möchte ich vor allem mein Badminton- und Faustballteam nennen, die nach der Arbeit schnell für Ablenkung vom Laboralltag gesorgt haben.

Ein großes Dankeschön geht an Martin Brune und Molly Strom für die besondere Unterstützung in diesem Studium und zahlreiche wissenschaftliche und private Diskussionen sowie unzählige Korrekturlesestunden meiner Abschlussarbeiten.

Von ganzem Herzen möchte ich mich bei meiner Familie bedanken, die in allen Phasen dieser Arbeit und des vorangegangenen Studiums eine große Unterstützung war.

Zu guter Letzt gilt mein größter Dank meinem Freund Philipp, der mir nicht nur während der Promotion auf unterschiedlichste Art und Weise eine sehr große Unterstützung war.

Veröffentlichungen im Rahmen der Dissertation

Feiner RC; Müller KM, Recent Progress in Protein-Protein Interaction Study for EGFR-Targeted Therapeutics. *Expert Rev. Proteomics 13* (9), 817–832 (2016) doi: 10.1080/14789450.2016.1212665.

Feiner RC, Teschner K, Schierbaum I, Teschner J, Müller KM, AAV production in suspension: evaluation of different cell culture media and scale-up potential. *BMC Proceedings* 12 (Suppl 1):P-349, (2018) doi: 10.1186/s12919-018-0097-x.

Feiner RC, Pennè I, Müller B, Müller KM, EGF-mCherry Fusion Protein Expressed in E. coli Shows Product Heterogeneity but a High Biological Activity. *Biochemistry* 58 (8), 1043-1047 (2019) doi: 10.1021/acs.biochem.9b00021.

Feiner RC, Teschner J, Teschner KE, Radukic MT, Baumann T, Hagen S, Hannappel Y, Biere N, Anselmetti D, Arndt KM, Müller KM, rAAV engineering for capsid-protein enzyme insertions and mosaicism reveals resilience to mutational, structural and thermal perturbations. Scientific Reports (2019) eingereicht.

Feiner RC, Tescher KE, Teschner J, Müller KM, HEK293-KARE1, a cell line with stably integrated adenovirus helper sequences simplifies rAAV production, BMC Biotechnology (2019), eingereicht.

Feiner RC, Kemker I, Krutzke L, Allmendinger E, Mandell DJ, Sewald N, Kochanek S, Müller KM, Computational design of EGFR-binding peptides enables tumor-targeting of recombinant adeno-associated virus capsids, eingereicht.

Gedruckt auf alterungsbeständigem Papier °° ISO 9706

Weitere Veröffentlichungen

Baier T, Kros D, Feiner RC, Lauersen KJ, Müller KM, Kruse O, Engineered Fusion Proteins for Efficient Protein Secretion and Purification of a Human Growth Factor from the Green Microalga Chlamydomonas reinhardtii. *ACS Synth Biol.* 7(11), 2547-2557 (2018) doi: 10.1021/acssynbio.8b00226.

Kemker I, Feiner RC, Müller KM, Sewald N, Size-dependent Cellular Uptake of RGD Peptides, *ChemBioChem*, https://doi.org/10.1002/cbic.201900512.

Kemker I, Feiner RC, Müller KM, Sewald N, Biocatalytic halogenation enables late-stage modification of RGD peptides, in Vorbereitung.

Gruß H, Feiner RC, Marion A, Jewgiński M; Kristian M. Müller, Latajka R, Sewald N, Peptide Stapling by Suzuki-Miyaura Cross-Coupling, in Vorbereitung.

Poster

Feiner RC, Schlicht K, Teschner K, Arndt KM, Müller KM, Recombinant Adeno-associated virus (rAAV) for tumor therapy: engineering of capsid and genetic modifications. 67. Mosbacher Kolloquium - "Protein Design: From First Principles to Biomedical Applications", Mosbach, 30.03.2016 - 02.04.2016.

Feiner RC, Teschner K, Schierbaum I, Teschner J, Müller KM, AAV production in suspension: Evaluation of different cell culture media and scale-up potential. 25th ESACT Meeting: Cell technologies for innovative therapies, Lausanne, 14-17.05.2017.

Feiner RC, Teschner K, Teschner J, Scheiner O, Müller KM, Recombinant adeno-associated virus for tumor therapy – capsid and genetic engineering. 4th Global Synthetic Biology & Gene Editing, London, 04.-05.12.2017.

Feiner RC, Müller KM, Re-targeting of recombinant adeno-associated viruses for enzyme mediated tumor therapy. XXIVth Annual Meeting DG-GT, Freiburg, 26-28.09.2018.

Contents

List of abbreviations

aa	amino acid
AAP	assembly-activating protein
AAV	adeno-associated virus
AAVR	AAV receptor
AFM	atomic force microscopy
APS	ammonium persulfate
bp	base pair
CHAPS	3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate
DAPI	4′,6-diamidino-2-phenylindole
DCM	dichloromethane
DIC	*N*,*N*-diisopropylcarbodiimide
DIPEA	*N*,*N*-diisopropylethylamine
DMF	*N*,*N*-dimethylformamide
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DTT	dithiothreitol
EDC	1-ethyl-3-(3-dimethylaminopropyl)carbodiimid
EDTA	ethylenediaminetetraacetic acid
EGF(R)	epidermal growth factor (receptor)
EGFP	enhanced green fluorescent protein
FACS	fluorescence activated cell sorting
FAM	carboxyfluorescein
FCS	fetal calf serum
gc	genomic copies
GFP	green fluorescent protein
HGFR	human growth factor receptor
HOBt	1-hydroxybenzotriazole
HRP	horseradish peroxidase
HSPG	heparan sulfate proteoglycane
IEX	ion exchange chromatography
IMAC	immobilized metal ion chromatography
ITR	inverted terminal repeat
LamR	37/67 kDa laminin receptor
LB	Luria-Bertani
LC/MS	liquid chromatography–mass spectrometry
MALDI	matrix-assisted laser desorption/ionization
NHS	*N*-hydroxysuccinimide
NMP	*N*-methyl-2-pyrrolidone

nt	nucleotide
OD	optical density
ORF	open reading frame
Oxyma	ethyl cyanohydroxyiminoacetate
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PDGFR	platelet derived growth factor receptor
PyAOP	azabenzotriazol-1-yloxy)tripyrrolidinophosphonium hexafluorophosphate
SIA	sialic acid
TEM	transmission electron microscopy
TEMED	*N*,*N*,*N*′,*N*′-tetramethylethane-1,2-diamine
TFA	trifluoroacetic acid
TFE	2,2,2-trifluoroethanol
TIPS	triisopropylsilane
VP	viral protein

Commonly used abbreviations and SI units are not separately listed.

1. Zusammenfassung

Rekombinante Adeno-assoziierte Viren (rAAV) haben herausragende Fähigkeiten in der Gentherapie und bieten somit die Möglichkeit der personalisierten Medizin. Sie zeichnen sich durch eine geringe Immunogenität, hohe Stabilität und langfristige Genexpression in Zielzellen aus. Um ihr Potenzial voll auszuschöpfen, kann die Manipulation auf genetischer Ebene genutzt werden, um gezielt bestimmte Zellen zu erreichen. Die Tumortherapie auf Basis dieser innovativen, neuartigen Plattformtechnologie bietet ein großes Potenzial für die Klinik. Die vorliegende Arbeit analysiert genetisch, mit Tumormarker-spezifischen Liganden, modifizierte rAAVs im Kontext verschiedener Serotypen und liefert zudem Einblicke in verschiedene rAAV-Produktionsoptionen.

Der epidermale Wachstumsfaktor-Rezeptor (EGFR), ein oft überexprimierter Marker in Tumorzellen, war der Zielrezeptor zweier rechnerisch entworfener Peptide (pep1jhf und pep1osy) und dem vom EGFR abgeleiteten Peptid EDA. Alle Peptide wurden chemisch synthetisiert, zyklisiert und mit 5(6)-Carboxyfluorescein markiert. Die sekundäre Strukturanalyse mittels Zirkulardichroismus zeigte eine deutliche β-Hairpin Struktur für eines dieser Peptide. Fluoreszenzpolarisationsexperimente demonstrierten eine geringe Affinität aller Peptide zum löslichen Rezeptor. Zur Bewertung deren Wirksamkeit wurden Zelllinien verwendet, die unterschiedlichen EGFR-Konzentrationen exprimieren. In einem weiterführenden Projekt wurde daher der EGF-Rezeptorzustand der Zelllinien mit Hilfe des Fusionsproteins EGF-mCherry charakterisiert. Ergebnisse aus Wundheilungsassays und der konfokalen Mikroskopie lebender Zellen passen gut zusammen und zeigten einen inhibitorischen Effekt auf die Zellproliferation und -migration sowie die spezifische Internalisierung eines Peptids in EGFR-überexprimierenden Zellen.

Zur spezifischen Bindung und Internalisierung von Zielzellen durch den AAV wurden darüber hinaus die drei Peptide genetisch in die Loopstruktur um den Aminosäurerest 587 des VP Proteins eingefügt. Die damit erzielte Präsentation von 60 Peptiden auf der Kapsidoberfläche führte zur Erhöhung der Avidität des Gesamtkomplexes. Als Folge dessen konnte die Fähigkeit zur spezifischen Transduktion von EGFR-überexprimierenden Zielzellen für zwei rAAV Varianten nachgewiesen werden. Erste Experimente an xenotransplantierten Chorioallantoismembranen (CAM) von Hühnerembryonen zeigten die Verteilung viraler genomischer Kopien in ausgewählten Organen und dem Tumorgewebe nach systemischer Injektion der rAAV-Peptidvarianten. Im Fall von rAAV2 pep1jhf gelang eine hohe Transduktion des Tumorgewebes, auch wenn die parallele Transduktion einiger Organe zu schweren Nebenwirkungen im Xenograft führte.

Das bestehende Plasmidsystem wurde erweitert und sollte das Einfügen unterschiedlicher Motive über genetische Modifikation, nicht nur in AAV2, sondern auch in anderen Serotypen ermöglichen. Zu diesem Zweck wurde das rAAV2-Plasmidsystem auf rAAV6 und rAAV9 übertragen und die

resultierenden viralen Vektoren charakterisiert. Begleitet wurden diese Studien durch die Optimierung des AAV-Produktionsprozesses in einer neuartigen Produktionszelllinie sowie Suspensionskultur.

Die Ergebnisse dieser Studien liefern wichtige Erkenntnisse für zukünftige Verbesserungen des Designs und der Anwendung von Peptid-modifizierten rAAV-Varianten. In Kombination mit transkriptionellen oder translationalen Targeting-Ansätzen kann so in Zukunft eine hochspezifische Therapie ermöglicht werden.

2. Abstract

Recombinant adeno-associated viruses (rAAV) combine outstanding gene therapy capabilities with excellent opportunities for personalized medicine. They provide low immunogenicity, high stability and long-term gene expression in target cells. To exploit their full potential, genetic manipulation can be used to enable targeting of specific cells. Tumor therapy based on this innovative emerging platform technology offers great potential for the clinic. This thesis analyzed tumor marker-based re-targeting of rAAVs in the context of different serotypes and besides provided insights into different rAAV production options.

The epidermal growth factor receptor (EGFR), which is an often overexpressed marker in tumor cells, was targeted by two computationally designed peptides (pep1jhf and pep1osy) and one EGFR derived peptide (EDA). Peptides were chemically synthesized, cyclized and labeled with 5(6)-carboxyfluorescein. Secondary structure analysis by circular dichroism showed significant β-hairpin structure for one peptide. In fluorescence polarization assays a low affinity was demonstrated. Cell lines expressing different levels of the EGF receptor were used to evaluate the efficacy of peptides. As s supplementary project the cell lines' EGF receptor state was characterized using the fusion protein EGF-mCherry. Wound-healing assays and live cell confocal microscopy fit well together and demonstrate an inhibitory effect on cell proliferation and migration as well as specific internalization of one peptide in EGFR-overexpressing cells.

Furthermore, the three different peptides were genetically inserted into the loop structure forming the three-fold spike of the AAV capsid at amino acid residue 587. By presentation of 60 peptides in the capsid surface the avidity of the complex is elevated. These rAAVs showed a transduction ability in combination with a high specificity for EGFR-overexpressing cell lines for two rAAV variants. First experiments in a whole organism were performed in egg xenotransplant chorioallantoic membrane (CAM) assays, where rAAV variants were injected systemically and distribution of viral genomic copies was determined in selected organs and the tumor tissue. Here, one variant showed high transduction efficiency of the tumor tissue and furthermore also of the organs, which led to severe side-effects in the xenograft.

The insertion of motifs by genetic construction of plasmids not only compatible with the production of AAV2, but also AAV6 and AAV9 was established. To this end, the rAAV2 plasmid system was transferred rAAV6 and rAAV9 and resulting viral vectors were characterized. These studies were accompanied by the optimization of the AAV production process in a novel producer cell line and in suspension culture.

Overall these studies provide important insights to improve design and application of peptide-modified rAAV variants in the future. Combining these results with transcriptional or translational targeting approaches may enable a highly specific therapy in the future.

3. Introduction

3.1. Gene therapy in the treatment of cancer

Cancer is the second leading cause of death worldwide.[1] The number of new cancer cases in Europe in 2018 was estimated to be 3.91 million. It was also estimated that 1.93 million people in Europe will die of cancer in 2018.[2] Research in this field is therefore of particular interest not only in Europe but all over the world. Common treatment of cancer includes surgery of the tumor tissue, chemotherapy and radiotherapy. However, these frequently used therapy methods are by no means applicable to all tumors. Therefore, new forms of therapy are constantly being researched. Gene therapy offers the possibility of a treatment specifically tailored to the patient, even for types of cancer that cannot be treated with conventional methods. The importance of gene therapy in cancer research is also reflected in the number of clinical trials: 65% of all gene therapy clinical trials were related to cancer.[3] The US Food and Drug administration (FDA) defined the term human gene therapy as a technique that modifies a person's genes to treat or cure disease.[4] Gene therapy in general does not only refer to the replacement of a dysfunctional gene with a healthy copy, but also includes inactivation of dysfunctional genes or introduction of new or modified genes to the cells. If this definition is now applied to the therapy of cancer, the following possibilities arise: (a) expression of a gene to induce apoptosis or increase tumor sensitivity to conventional drug/radiotherapy; (b) employing a wild-type tumor suppressor gene to compensate for its loss/deregulation; (c) blocking expression of an oncogene using an antisense approach (RNA/DNA); and (d) enhancing immunogenicity of the tumor to stimulate recognition of immune cells.[5]

All of these strategies require the transport of the gene into the target cells. Basically, a classification is made between viral and non-viral strategies.[6] Transfection of non-viral vectors into target cells require mostly chemical or biophysical methods. Nucleic acids can be introduced into the cell via nanoparticles, liposomes or electroporation.[6–8] Production of non-viral vectors is often not limited and thus cost-effective. However, these methods have the disadvantage over viral vectors that they can enter the target cells less efficiently. In nature, viruses transport their genome for replication in host cells. This process is generally referred to as transduction. In several viral systems it is possible to exchange the viral genome for an artificial promoter and a gene of interest. Since viral vectors have evolved their cell entry and DNA release mechanism over decades, these systems are more challenging for introduction of modifications compared to non-viral systems. Nevertheless, more than 67% of gene therapy clinical trials have been performed using viral vector systems.[3] These were mainly based on work with adenovirus, retrovirus, adeno-associated virus and lentivirus. Most of these viruses trigger an immune response in the human body. The adeno-associated virus (AAV) is characterized by very low pathogenicity and also has the advantage of high transduction efficacy and long-term gene expression in the target cell.[9] The past showed that working with AAV is promising as first therapeutics were already on the market. The latest agent, *voretigene*

neparvovec (Luxturna), was approved by FDA and EMA in 2018 and is based on rAAV2.[10] Luxturna delivers a functional variant of the RPE65 gene to supplement biallelic RPE65 mutation-associated retinal dystrophy.

3.2. Biology of adeno-associated viruses

AAV is classified as a non-enveloped viral vector belonging to the family of *parvoviridae*. The single-stranded DNA genome of AAV has a size of only 4.7 kb and is flanked by inverted-terminal repeat sequences (ITRs). The left main open reading frame (ORF) provides genetic information for four non-structural Rep proteins, that have been shown to play a major role in replication, packaging and genomic integration.[11,12] Expression of these proteins is driven by two viral promoters p5 and p19. The right ORF encodes the structural VP proteins (VP1, VP2, VP3) that assemble the viral capsid in a 1:1:10 ratio. Expression of all VP proteins is driven by the p40 promoter. Proteins with an identical C-terminus but differing N-termini derive from alternative splicing and leaky-scanning mechanisms (Figure 1).[13,14] Another reading frame within the *cap* genes was recognized, which encodes the non-structural assembly activating protein (AAP) from a non-canonical CTG start codon.[15,16] Recently it was shown that AAPs influence on capsid assembly is dependent on the serotype.[17] Especially AAV8, AAV7 and AAV2 were shown to be strongly dependent in contrast to AAV3 and AAV9. In the same reading frame, the sequence of the x protein was found, which supposedly is involved in AAVs DNA replication.[18] Besides these nested reading frames, it has been shown previously that in addition, other proteins are encoded whose functions are not yet known.[14]

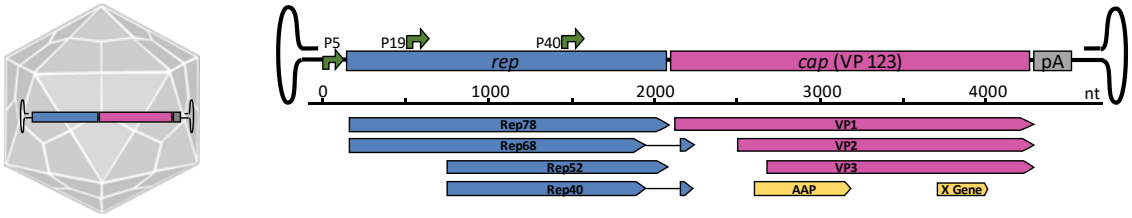


Figure 1: Schematic representation of the rAAV genomic structure. The genome with a size of 4.7 kb is encapsidated in the capsid shell (left). Three reading frames highlighted in blue, pink and orange are under control of the promoter p5, p19 and p40. Besides the non-structural Rep proteins (blue) and the structural VP proteins (pink), the genome codes for other proteins as AAP and x (orange) which are necessary for replication and capsid assembly.

A total of 60 VP proteins is required to build a capsid shell with T=1 icosahedral symmetry. The structure of an AAV2 was solved with 3 Å resolution using X-ray crystallography.[19] To date, further structures of other serotypes have also been resolved, e.g. AAV6[20], AAV9[21] and it was shown that all serotypes share a high structural homology. Comparison of VP3 proteins from different serotypes show highly conserved regions (Figure 2A). The core of the VP3 proteins consists of an eight-stranded β-barrel motif and a small α-helix.[22] Nine variable regions (VR-I to VR-IX) have been described which can be found on the surface of the assembled capsid and which differ largely between the different serotypes (Figure 2B).[21]

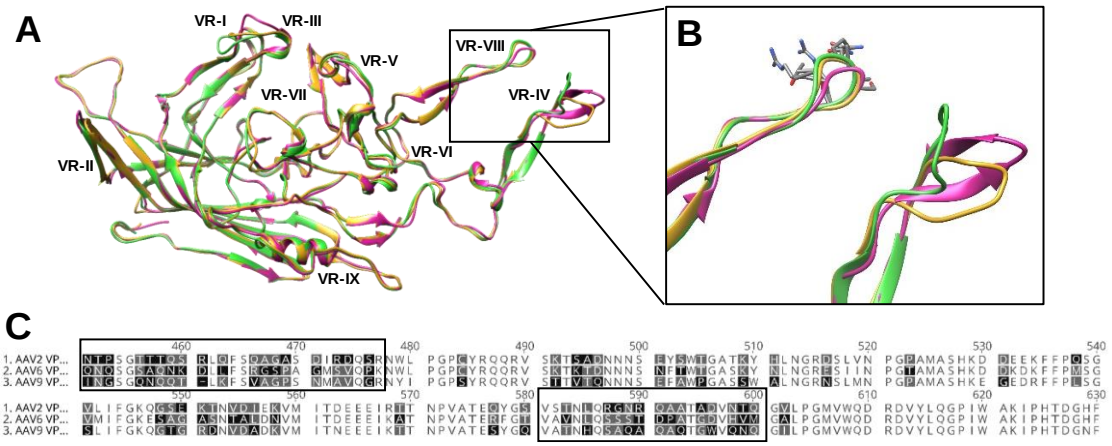


Figure 2: Alignment of VP protein sequences from AAV serotype 2, 6 and 9. (A) Structural alignment of crystal structures using UCSF Chimera. (AAV2 PDB 1LP3 pink, AAV6 PDB 3OAH orange, AAV9 PDB 3UX1 green). The three serotypes show great identity. Differences can be observed for some regions, which were named variable region VR-IV and VR-VIII. (B) A great variance between different serotypes is observed in the variable region IV and VIII, which are further related to as 453- and 587- loop region. Arginine residues of AAV2 (R585 and R588) are indicated with atoms in VR-VIII. (C) Amino acid alignment of the three serotypes (NCBI Reference sequence: NC_001401.2; Genbank: AAV6 AF028704.1; AAV9 AY530579.1). Differences in amino acids sequences are highlighted in grey shades. Variable region VR-IV and VR-VIII are bordered to indicated high sequence variation.

AAV capsids of all known serotypes share two-, three- and five-fold symmetry axes that determine the characteristic appearance of the capsid shell with depressions and protrusions. Between different serotypes some regions remain highly conserved, while others are varying. An extraction of the alignment of serotypes 2,6 and 9 is shown in Figure 2C and highlights the differences in VR-IV and VR-VIII. Variable regions IV and VIII, mainly located in the three-fold spike, determine the serotypes tropism for a specific receptor interaction (see Table 1). Different receptors are known to induce interactions with the host cell, but until now, not every detail is fully covered.

Table 1: Summary of the broad diversity in cellular receptors that are necessary for AAV transduction of different serotypes.

Serotype	Primary receptor	Coreceptor
AAV1	α2-3/α2-6 N-linked sialic acid (SIA)[23]	
AAV2	Heparan sulfate proteoglycan (HSPG)[24]	fibroblast growth factor 1 (FGFR-1)[25] hepatocyte growth factor (HGFR)[26] αVβ5 integrin[27] α5β1 integrin[28] 37/67 kDa laminin receptor (LamR)[29]
AAV3	HSPG[30]	FGFR-1[31] human HGFR[32] LamR[29]
AAV4	α2-3 *O*-linked SIA[33]	
AAV5	α2-3 *N*-linked SIA[33,34]	Platelet-derived growth factor receptor (PDGFR)[35]
AAV6	HSPG[36] α2-3/α2-6 *N*-linked SIA[23]	Epidermal growth factor receptor (EGFR)[29]
AAV8		LamR[29]
AAV9	Terminal *N*-linked galactose[37]	LamR[29]

Serotype AAV2 was studied extensively in the past and amino acids contributing to receptor binding where identified in the three-fold spike. The basic residues R484, R487, K532, R585 and R588 are necessary for interaction with the primary receptor heparan sulfate proteoglycane (HSPG).[24,38,39]

Further receptor-mediated clathrin-dependent endocytosis is mediated by interaction with other receptors as fibroblast growth factor 1 (FGFR-1)[25], hepatocyte growth factor (HGFR)[26], 37/67 kDa laminin receptor (LamR)[29] and integrins (α5β1 and αvβ5)[27,28]. Recently, a gene product of KIAA0319L was discovered to be a multiserotype receptor which was designated AAV receptor (AAVR).[40] It was shown that most serotypes, except AAV4, are dependent on this protein.[41] Until now it is not fully understood which role this potential receptor is playing in the cellular uptake of viral particles.

After attachment and host cell entry the viral vector is transported into the cells via receptor-mediated endocytosis from clathrin-coated vesicles.[42] Upon endosomal escape, AAV is able to enter the nucleus of the host cells and the ssDNA genome is released. AAVs have the unique ability to integrate into the genome of the host cell. The integration of the AAV genome is performed specifically at the AAV safe-harbor locus (AAVS1) in chromosome 19 when no helper virus is present and the AAV enters the lysogenic cycle.[43,44] Different parts of the wild-type AAV play a crucial role in genomic integration, e.g. ITRs and Rep78 and Rep68.[45] If the host cell is additionally infected with a helper virus, such as herpes simplex virus (HSV), the AAV enters the lytic cycle. This leads to subsequent DNA replication and gene expression of the viral proteins. The capsid assembles from VP proteins in the nucleoli of the cells before single-stranded DNA is packaged.[15,46]

Recombinant AAVs (rAAVs) are derived from the natural system and are intended to introduce transgenes specifically into target cells. The *rep* and *cap* genes of the wild-type AAV are therefore replaced by an expression cassette containing a gene of interest under the control of a suitable promoter. During the exchange of the expression cassette the ITR sequences need to stay intact. Both sequences are crucial for packaging of the genome into the assembled capsid. Production of recombinant AAVs is only possible if *rep* and *cap* gene are supplied *in trans* on another plasmid. In the past, rAAVs were produced by a combined system of plasmid transfection and infection with an adenovirus. However, this production variant has the major disadvantage of contamination of the final vector preparation by the adenovirus. Another disadvantage is that higher safety levels are required due to the introduction of a pathogenic virus. Major breakthrough came at the end of the 1990s when two independent groups were able to demonstrate that rAAV production in HEK293 cells is possible via a helper plasmid.[47,48] The helper plasmid provides the necessary adenoviral elements E2A, E4 and VA RNA and the HEK293 cells line provides still missing E1A and E1B. Using this method, replication-deficient rAAV preparations, which are wild-type- and adenovirus-free, can be obtained.

3.3. Improving AAV vectors for tumor therapy

Gene therapy using rAAVs was shown to be promising also in the field of cancer research. Therapy of cancer with viral vectors requires a selective transduction of the tumor cells. The natural serotypes of AAVs offer the ability to transduce specific tissues via their unique tropism. A problem in

therapy with natural serotypes is the high immunoprevalence of the population.[9] Studies show that half of the population has already come into contact with AAV2 and has thus been able to produce anti-AAV2 neutralizing antibodies.[49] Another problem in the treatment of cancer is the biodistribution in tissue as well as the targeted transduction of tumor cells. Overcoming of these strong limitations was pursued with various approaches.

Delivery and entry of target cells is induced via the amino acid sequence that defines the capsid shell. Solving the crystal structure provided necessary information on the capsid surface and the exposure of amino acids to the surrounding. Two general methods were applied to change AAVs tropism: rational design of targeting motifs and directed evolution of the capsid proteins.

Directed evolution of AAVs capsid proteins does not require a good knowledge of the capsid structure and transduction mechanisms. In 2003, Müller et al. showed the feasibility of an *in vitro* random peptide library system for AAV2.[50] An approach by Michelfelder et al. also used a random peptide library displayed on the viral capsid in *in vitro*.[51] The selected library-derived rAAVs transduced tumors *in vivo*.

For rational engineering of the capsid, it must be known at which surface-exposed positions of the VP protein integrations can be carried out without loss of productivity. In the past several groups determined possible insertion sites. For AAV2 two groups have demonstrated that integration of peptide ligands in VP proteins at residue positions 46, 115, 139, 161, 261, 381, 447, 459, 534, 573, 584, 587, and 588 did not interfere with capsid assembly (Figure 3A).[52,53] The targeting peptide to be integrated into the capsid should be structure-independent and not too large to avoid destabilization of the capsid shell.[54] Retargeting towards a new target molecule requires the neutralization of the natural tropism. Main interaction between viral particles and primary receptors was described already for a few serotypes. As the best characterized serotype AAV2 all amino acids interacting with HSPG are well known and mutations in the two arginine residues R585 and R588 allow for strong reduction of transduction efficiency.[55] For AAV6 both residues are missing and thus not contributing to the interaction with HSPG. Here a lysine residue mediates the interaction and it was shown that a K531E mutation impairs affinity towards HSPG.[56] The amino acids inducing the interaction with the secondary receptor EGFR are not characterized yet. The primary receptor N-linked galactose in AAV9 is strongly linked to two amino acids N272 and W503 that are known to be important for binding.[37]

Integration into the capsid can be directly or indirectly targeting a tumor cell specific feature. In direct approaches a peptide ligand is able to bind a cell-specific target, while in indirect targeting approaches, the interaction with the target cell is mediated via an associated molecule, which is bound to the capsid surface. Indirect targeting was described in the past for an rAAV2 that present a minimal immunoglobin G (IgG) binding domain Z34C in amino acid position 587 (Figure 3B).[57] rAAVs were loaded with different antibodies and specific transduction of human hematopoietic

cell lines was observed. Integration of motifs also allows for site-specific bioorthogonal labelling of rAAV particles. Previous work showed that integration of the recognition motif for the formylglycine-generating enzyme (FGE) was possible at amino acid position 587 and allowed for covalent conjugation of the resulting aldehyde either with Alexa488 hydrazide or amine-functionalized gold particles (Figure 3B).[58]

Direct targeting for rAAV2 was shown with different peptide integrations in position 587 but also with VP2 *N*-terminal fusions. RGD peptides have been incorporated into surface-exposed VP areas.[53] It was shown that cells were transduced independently from the natural HSPG motif. The *N*-terminal fusion of even whole proteins to the VP2 protein was demonstrated by different groups.[59,60] In both approaches a four-plasmid system is required where the VP2 fusion protein is delivered separately from VP1 and VP3 (Figure 3C).

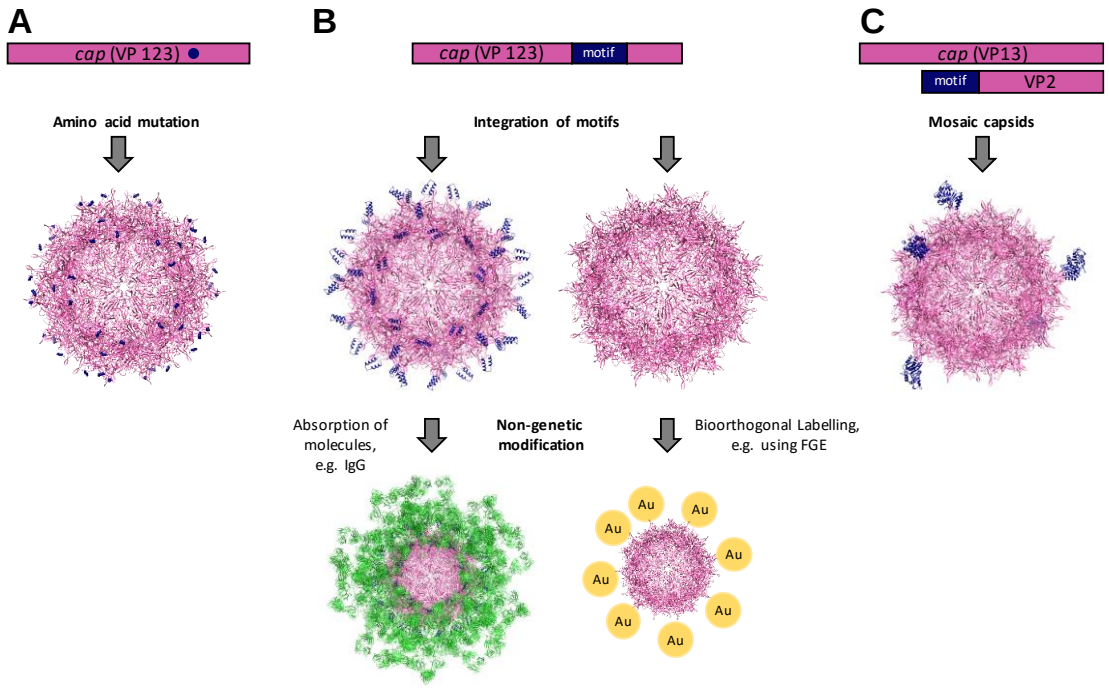


Figure 3: Strategies for rAAV retargeting. (A) Single amino acid mutations can be genetically introduced. (B) Site-specific integration is tolerated at various sites of the VP proteins and enables for further non-genetic modifications of the capsid. Absorption of IgG molecules to integrated Z34C domains was shown previously to results in biologically active vectors. Integration of motifs also allows for further biorthogonal labelling, e.g. using the formylglycine generating enzyme (FGE) to generate an aldehyde available for covalent conjugation of amine-functionalized gold particles (Au). (C) Integration of motifs or fusion proteins does not necessarily result in fully-modified capsids. Generation of so-called mosaic viral vectors remains possible with an alteration in the plasmid system.[59,60]

All systems described were developed for serotype 2 but since the capsid similarity between serotypes is high, some groups also established incorporation of peptides into other serotypes. From random peptide libraries selected peptides have been transferred from serotype 2 to serotype 8 and 9.[61] Here it was shown, that not only the peptide sequence optimized for AAV2 determines the transduction ability *in vivo* but also the overall capsid contributes to the tropism. A different approach relying on retargeting using RGD peptides was shown in AAV6.[62]

In addition to the specific targeting of the cell by the viral capsid, tumor therapy can also use other properties that e.g. antibodies do not have. The viral particles do not introduce any active substance into the cells. Instead, only the DNA is specifically delivered under the control of a promoter. The expression of the target protein can be specifically activated in tumor cells in various ways. One example is the promoter of the C-X-C chemokine receptor type 4 (CXCR4). For this promoter it had been shown in the context of AAV2 that expression of the transgene was only achieved in tumor cells.[63] A second approach is the use of anti microRNAs (miRNA). The Let-7 family has twelve known members that target the same mRNA sequenes.[64] Members of this family are known to accumulate in differentiated cells, but were shown to be downregulated in cancer cells by mechanism that are not fully understood.[65] After transduction into healthy cells, the mRNA of the delivered transgene is degraded by binding the anti-Let-7 miRNA, while translation can take place in tumor cells.[66] Specific targeting combining both approaches can be summarized under the term virus-directed enzyme prodrug therapy (VDEPT).[67] This emerging strategy in the treatment of cancer allows for direct targeting of cells via tumor specific features. By delivering of an enzyme, a prodrug is activated into a cytotoxic compound that finally leads to apoptosis of the cancer cell. In the past, several targets of cancer cells have been identified and used for therapy. The epidermal growth factor receptor (EGFR) is one example for a validated tumor target.[68]

3.4. Targeting the epidermal growth factor receptor

The EGFR (ErbB1) is a member of the ErbB family and known to be ubiquitously expressed in epithelial, mesenchymal and neuronal cells and their cellular progenitors.[69] The receptor can be subdivided into an extracellular, a transmembrane and an intracellular domain (Figure 4). The extracellular domain is able to bind up to seven different ligands and induces subsequent dimerization with another member of the ErbB family. The dimerization induces intracellular tyrosine kinase activity and further phosphorylation. Several different cell signaling pathways are known to be activated upon phosphorylation, which are involved mainly in cell growth and migration as well as proliferation and differentiation. The cellular response varies not only with the type of bound ligand but also on the type of receptor dimer pair.[70,71] Contributing to this variety of signaling pathways, is it not surprising that the EGFR is involved in several different types of cancers. Thus, the receptor, which is often overexpressed and causes aberrant signaling, is a validated target in cancer therapy.[72,73]

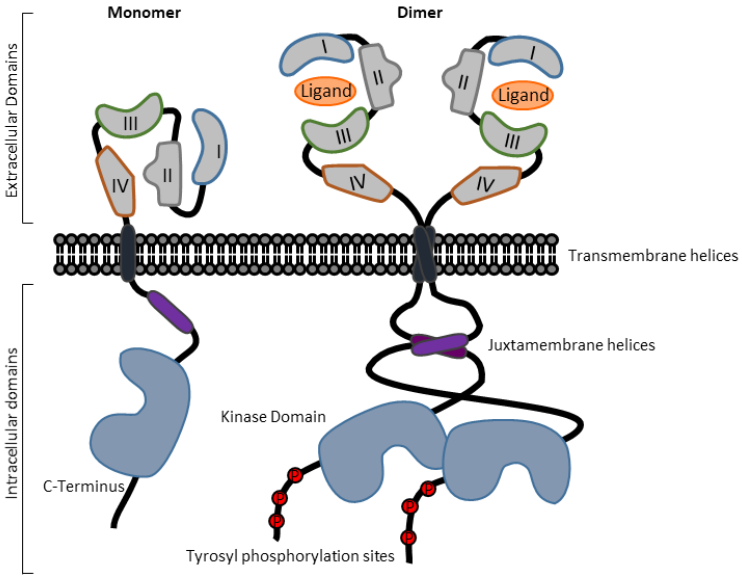


Figure 4: Schematic representation of the EGFR structure as a monomer (left) and as a dimer (right). The N-terminal extracellular domain is build up from four subunits I-IV followed by a transmembrane helix, a juxtamembrane segment and the kinase domain. (a) As a monomer without ligand the receptor is mainly present in the autoinhibited form in which domain II and IV are tethered by the dimerization arm. (b) Natural ligands are bound in an open conformation sandwiched between domain I and III, whereby the dimerization arm is exposed promoting dimerization. After homo- or heterodimerization internalization starts and the intracellular kinase domains form an asymmetric complex initiating phosphorylation. (This figure is part of the review article "Recent progress in protein-protein interaction study for EGFR-targeted therapeutics." which was published during this work).[74]

In the early 1980s overexpression of the EGFR was described by several groups of researchers. Shortly afterwards, different groups started targeting the receptor and blocking the ligand binding site. Initial experiments proved the hypothesis that monoclonal antibodies are able to interrupt receptor dimerization and subsequent signaling processes to prevent cell proliferation.[75] Ever since, researchers have been looking for EGFR antagonists that inhibit cell signaling. A detailed description about EGFR extracellular targeting approaches was summarized in a review article that was written during this work.[74]

Monoclonal antibodies (mAbs) represent the first group of EGFR biologic antagonists that are already FDA and EMA approved to cancer therapy. The first mAb brought to the clinic was C225 (Cetuximab), which is now marketed under the name Erbitux.[76] The epitope for EGFR-binding was characterized by co-crystallization and is shown in Figure 5A-C. As the natural ligand EGF, C225 binds to subunit III of the extracellular domain and competes with EGF binding. The antibodies panituzumab (Vectibix) and necitumumab (Portrazza) were approved in 2006 and 2015, respectively. A main problem of therapeutic antibodies lies within their properties, because biodistribution and the ability to penetrate tumor tissues is greatly reduced for large molecules.[77,78] Much smaller proteins that are able to bind specific targets are synthetically designed molecules as designed ankyrin repeat proteins (DARPins). Specific binding interactions are possible by residue randomization and ribosome or phage display selection.[79] Using these method, the four EGFR-binding DARPins E01, E67, E68 and E69 were identified and characterized.[80] Binding of DARPin

variants was described and it was shown that E01 and E68 both bind to subunit III of the extracellular domain, while E69 binds to domain I (Figure 5D,E)

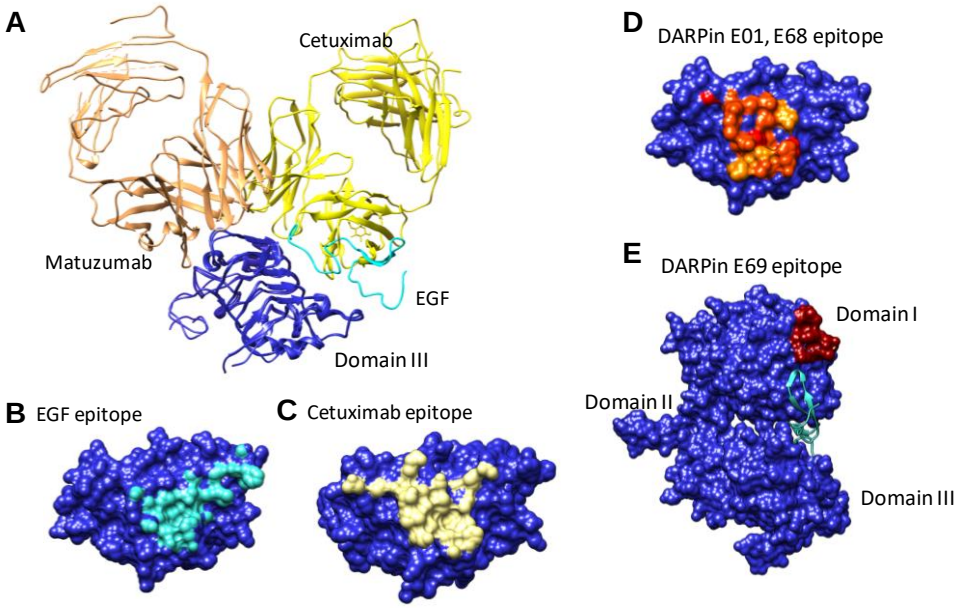


Figure 5: Crystal structure analysis of protein-protein interactions with domain III of the EGFR. (A) The position relative to domain III (blue) for the EGFR antagonists Matuzumab (brown) and Cetuximab (yellow) is shown in comparison to the natural ligand EGF (cyan). For reasons of simplification only the Fab fragments are shown in combination with domain III. (B, C) The epitope of EGF (cyan) and Cetuximab (yellow) are mapped on domain III. Structures were created from PDB files: EGFR [1IVO], Cetuximab [1YY9]. (D) DARPin E01 (red) and E68 (yellow) epitopes are mapped on domain III based on mutational analyses by Boersma et al..[80] *The merge area shows the overlap of both epitopes (orange). (E) DARPin E69 (red) binds to domain I of the EGFR as confirmed by mutation analysis experiments.*[80] *(This figure was modified from the review article "Recent progress in protein-protein interaction study for EGFR-targeted therapeutics." which was published during this work).*[74]

Even smaller targeting molecules can be found in peptides. They are able to diffuse into tissues and furthermore show a low immunogenicity. An overall advantage is the relatively low cost for production in comparison especially to antibodies. The best characterized synthetic EGFR-binding peptide was named GE11 (YHWYGYTPQNVI).[81] The linear dodecapeptide was selected from a phage-display library with the extracellular domain of EGFR. For the radioactively-labelled 12-residue peptide a dissociation constant of 22 nM was measured in cellular assays and internalization into EGFR-overexpressing cells was described.[82] However, surface plasmon resonance assays with purified human EGFR resulted in a dissociation constant of 459 µM.[83] A peptide found by computer assisted design (CAD) using EGFR crystal structures was named D4. The 6-mer linear peptide (LARLLT) shows significant binding to the receptor in cell culture experiments and also for a cyclic variant of the peptide activity was demonstrated.[84] Besides these peptides other approaches are more relying on the EGFR structure, especially domain II which is inducing the dimerization of two receptor molecules. Different approaches focused on targeting domain II and found the minimal binding motif of the so-called EGFR dimerization arm (QTPYYMNT). Different groups proved that this peptide is able to interact with the EGFR and provides inhibitory effects.[85,86]

Peptides are able to diffuse into tissues without interaction with a specific receptor. Affinity towards receptors is known to be lower for peptides in comparison to antibodies and DARPins. Conjugation to larger polymers as polyethylenglycol (PEG) was described previously to create stable molecules.[87] Multiplying the affinities of peptides in more complex molecules offers the opportunity to convert a low-affinity ligand into a high-avidity molecule. For the minimal EDA peptide, the generation of a bivalent molecule was already published in 2018.[88] Molecules combining an even larger number of low-affinity ligands can be chemically synthesized as nanoparticles. Here a broad range of liposomes and even integration of the GE11 peptide into a filamentous plant virus-based nanoparticles had been described.[89–92] This last approach relied on a system already existing in nature. Here, a potato virus X (PVX) served as a platform for incorporation of the GE11 peptide. Incorporation of peptides or peptide libraries into rAAVs was also shown in the past and offers the opportunity to present peptides in all 60 capsid proteins composing the capsid.[53,59,93,94] Providing peptides on such a megadalton complex might increase avidity and prohibit unwanted diffusion into cells, thus enhancing specificity of the targeting process.

4. Aim

Adeno-associated viruses are versatile carriers of DNA that offer a wide range of applications in the field of gene therapy and are also of increasing importance in tumor therapy. The natural tropism of AAVs does not allow for a specific transduction of target tissues. Ligand-mediated transduction offers the opportunity to increase specificity in target cell transduction and furthermore enables for de-targeting of the AAVs natural tropism. Approaches to direct the AAV towards the well-known tumor marker EGFR have been made in the past using incorporation of larger EGFR-binding proteins.

The fundamental question of this study was, whether computationally designed peptides are able to target the EGF receptor and allow for specific transduction of EGFR-overexpressing cells in context of the adeno-associated virus. In order to answer this question, synthetic peptides should be examined for their affinity to the receptor and their properties in cell culture experiments.

Furthermore, analysis of peptides in the recombinant AAV system requires production and purification of the viral particles. The increase in the quality of the virus preparations is another fundamental issue. Can the production and the purification method be improved in such a way that the previous work in this field can be substantially improved? To this end, effort should be made in the design of AAV production cell lines and plasmids systems as well as purification methods.

Ultimately, final rAAV variants will be analyzed with regard to the transduction of the target cells. The EGF receptor status of the cell lines to be investigated is important to detect the efficacy of the re-targeted rAAV variants. Therefore, the question arose how this status can be detected in a straightforward and convenient way.

4 Aim

5. Results and Discussion

5.1. Development of a EGF-mCherry fusion protein

Results of this projects were published with the title 'EGF-mCherry fusion protein expressed in *E. coli* shows product heterogeneity but high biological activity' in Biochemistry and are summarized in this chapter. The original publication is included in the appendix.

Targeting of tumor cells is a main task in therapy and relies on knowledge about the cells receptor state. Determination of EGFR receptor levels, internalization and inhibition of ligands was mostly analyzed using [^{125}I]-labelled EGF. Problems can be found in the short half-life of the isotope as well as the need for special equipped laboratories. Using fluorescent proteins instead is environmentally friendly and compatible with most modern workflows. Creation of the human EGF with *N*-terminal attachment of an EGFP was shown previously.[95] Further development was now carried out to make this fusion protein useful for a broader range of applications. From crystal structure analysis and the known cell bound state of full EGF we hypothesized that fusion to the *C*-terminus of mature EGF is equally if not better suited for fusions because it is more distant to the necessary disulfide bridges (Figure 6A,B).

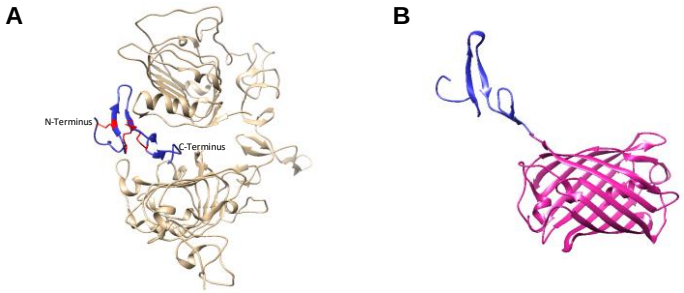


Figure 6: Crystal structure analysis of EGF and EGF-mCherry. (A) EGF:EGFR crystal structure proving both termini to be flexible for fusion to other proteins (PDB ID 1IVO). Three essential disulfide bonds for EGF a highlighted in red. (B) Assembly of EGF-mCherry from PDB ID 1JL9 and 2H5Q. (This figure was adopted from the supporting information of the article "EGF-mCherry fusion protein expressed in E. coli shows product heterogeneity but a high biological activity." which was published during this work).

5.1.1. Biochemical and biophysical characterization of EGF-mCherry

A fusion of the human EGF with mCherry attached to the *C*-terminus was generated in a pET21a vector using standard molecular-biological cloning procedures. Expression in *E. coli* BL21(DE3) with subsequent purification resulted in good yields of 14.9 mg per 500 ml shaker culture in LB medium. SDS-PAGE analysis under reducing and non-reducing conditions revealed that EGF-mCherry was mainly found in the cytosolic fraction and that partial dimer formation takes place, which can be reduced upon DTT addition (Figure 7A, B). Further analysis of MS and SDS-PAGE data revealed fragmentation at the chromophores *N*-acylimine at elevated temperatures.[96] Disulfide bond formation is crucial for biological function, because three internal disulfide bonds need to be formed. Due to the reductive nature of the cytosol, disulfide bonds cannot be formed straight during production[97]. We hypothesize that formation takes places upon cell disruption and purification by

exposure to the oxidative environment. For mass spectrometry using an LC-ESI-TOF a reduced and non-reduced sample of the final product were analyzed (Figure 7C). Each disulfide bond formation results in a mass difference of 2 Da in the spectrum. For the purified product a shift of 6 Da is expected as three disulfide bridges should be reduced to thiols. A shift of approximately 6.7 Da is observed for EGF-mCherry leading to the assumption that the proteins folding is correct.

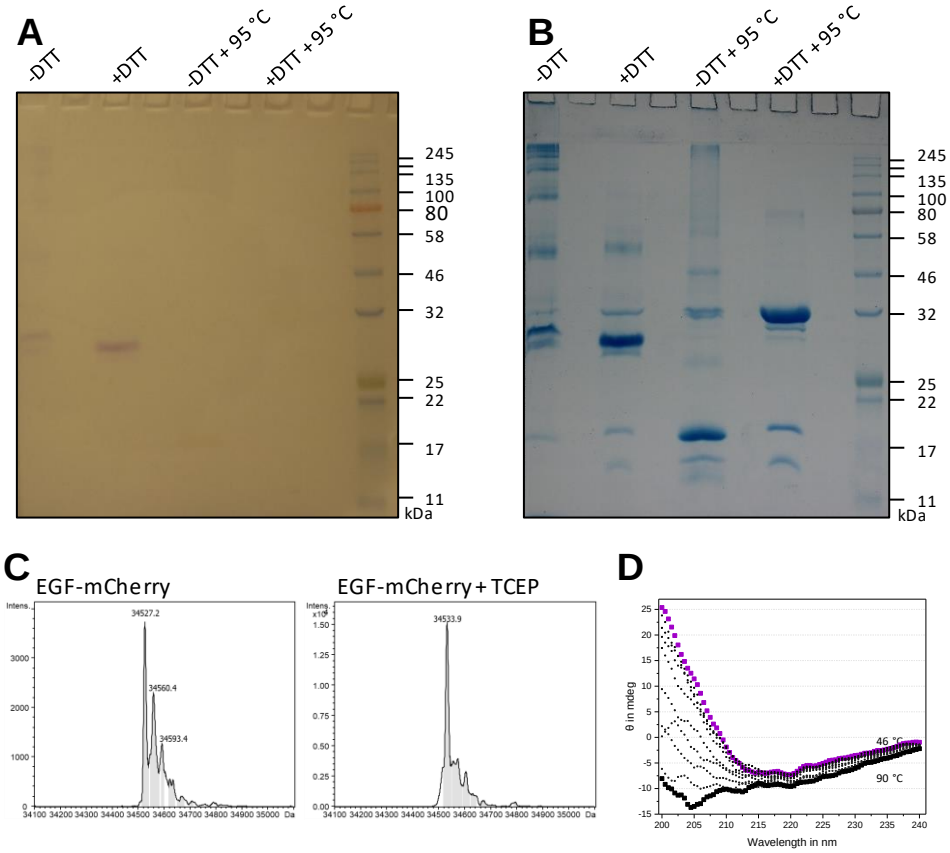


Figure 7: Characterization of the fusion protein EGF-mCherry. (A, B) SDS-PAGE of EGF-mCherry under varying conditions. EGF-mCherry was incubated with reducing (+DTT) and without reducing agent (–DTT) either at room temperature or at 95 °C for 10 min before application to a 12% SDS-PAGE. Images under white light are shown before (A) and after staining (B) with Coomassie Blue. (C) LC-ESI-MS analysis of EGF-mCherry under non-reducing and reducing conditions. (D) Circular dichroism spectra of a temperature course experiment from 46 to 90 °C in 1 °C steps.

Analysis of structural properties and thermal stability was assayed using circular dichroism spectroscopy. Even at elevated temperatures from 46-90 °C structural elements were detected in the sample, which is in accordance with literature values for mCherry.[98] Since of biophysical properties of EGF are in correspondence with data from literature, the affinity to the receptor and the biological activity must now be confirmed.

5.1.2. EGF-mCherry shows nanomolar binding affinities

Assaying the biological function of EGF-mCherry was carried out mainly in cell culture experiments. First experiments on cells expressing varying levels of EGFR on their cell surface revealed that only nanomolar concentrations are needed to detect a meaningful signal in microscopy, flow cytometry and microplate reader experiments. In wound healing experiments the biologic activity

of the EGF derivative was assayed on A431 cells. Cell migration was analyzed over the course of the experiment and closure of the gap was measured using ImageJ (Figure 8A). Cell migration rates were calculated from linear regressions as seen in Figure 8B. Comparing EGF-mCherry with the commercially available hEGF (Gibco), elevated migration rates were observed proving the biologic activity of EGF in the EGF-mCherry fusion construct Figure 8C. This assay was established to analyze EGF activity and was also performed in collaboration with the Algae Biotechnology & Bioenergy Research Group, Bielefeld.[99]

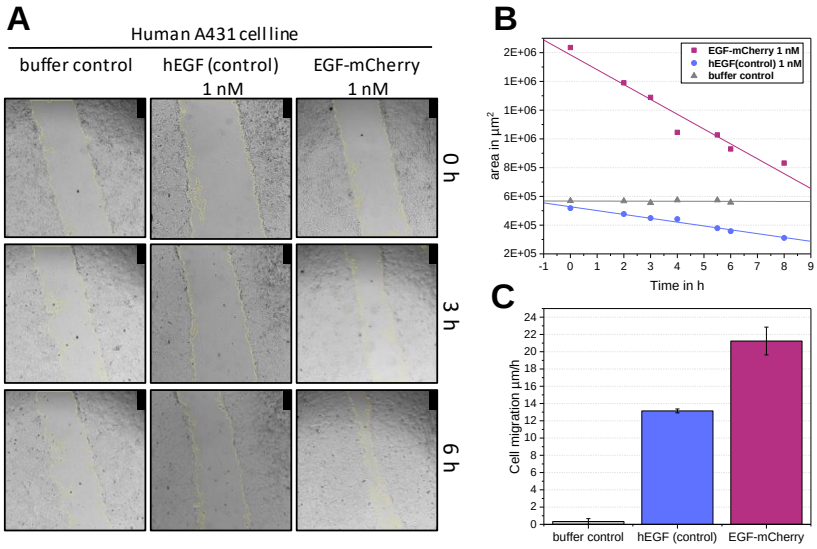


Figure 8: Wound healing assay for EGF-mCherry on A431 cells. (A) Cells were incubated with 0 nM (buffer control) and 1 nM EGF-mCherry or a commercial hEGF over a time course of 8 h. Cell-free areas and cell migration rates were determined using the ImageJ software plugin MRI Wound Healing Tool. Scale bars in upper right corner represent 250 µm. (B) The area of the closing wound was plotted against the time. Linear regression using Origin2019 were performed and cell migration rates calculated form the slope. (C) Bar chart of cell migration rates of the buffer control, the commercial hEGF (Gibco) and EGF-mCherry.

Further work in cell culture experiments was performed to visualize EGF receptor binding and internalization. The cell lines A431, Hela and MCF7, presenting decreasing levels of receptor, were analyzed using microscopy after incubation with 5 nM EGF-mCherry. Analysis revealed that the decrease in receptor level corresponds to a decrease in fluorescence intensity (Figure 9A).[100,101] Flow cytometry allowed a more detailed analysis and thus, the EGF receptor state was characterized for several different cell lines Figure 9B. These values are in good correspondence with values described in literature and allow for assessment of EGFR density of uncharacterized cells lines, e.g. HEK293 EGFR. Time-dependent microscopy images demonstrated the biologic activity on A431 cells via their internalization. Figure 9C shows microscopy images of A431 incubated with EGF-mCherry after 15 min and after 30 min in comparison to a control. Internalization of the fusion protein can be visualized and indicates that the mCherry fusion does not inhibit the function of EGF. Not only cell culture experiments were used to prove the EGFR binding affinity. Biolayer interferometry (BLI) experiments were used to characterize EGF-mCherry:EGFR binding affinity. The soluble domain of the EGFR (sEGFR, residue 1-621) was expressed in 293F cells, purified via

IMAC and subsequently immobilized on an amine-reactive biosensor (ARG2). Different concentrations of EGF-mCherry and as negative control bovine serum albumin (BSA) were measured using the BLItz system. Obtained curves showed a specific binding of EGF-mCherry to the sEGFR (Figure 9D) and a global fit based on a 1:1 kinetic binding model yielded a dissociation constant K_D of 7.6 nM. This value is in good agreement with the K_D of 1.8 nM reported for hEGF in surface plasmon resonance experiments.[83]

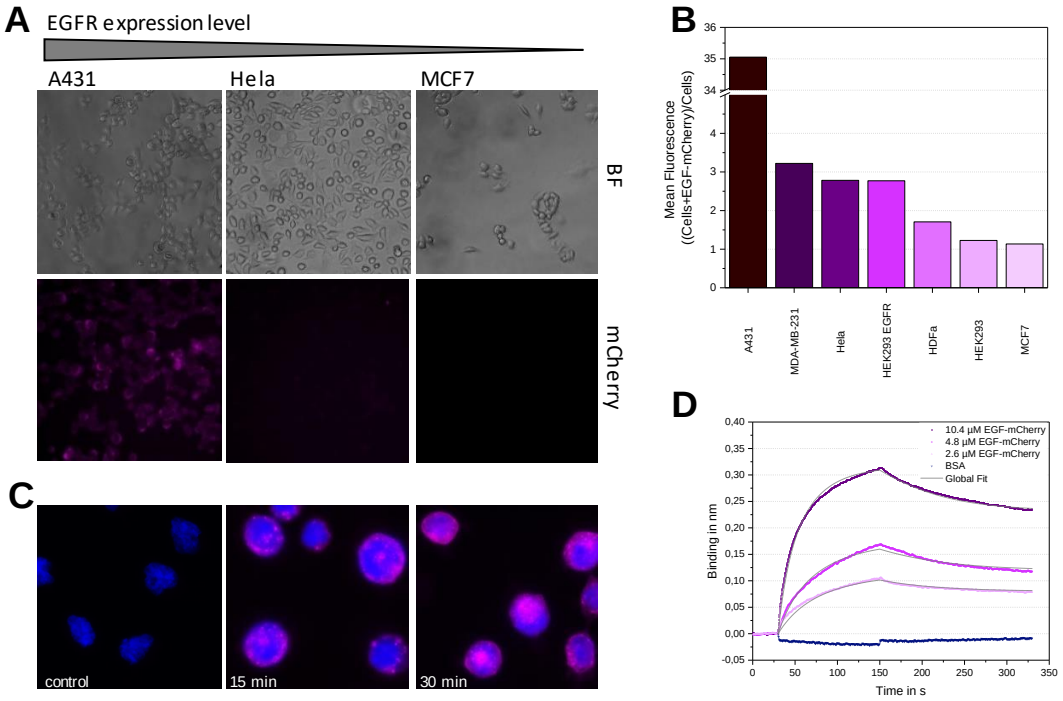


Figure 9: Characterization of EGF-mCherry binding affinity. (A) A431, Hela and MCF7 cells were incubated with 5 nM EGF-mCherry for 15 min before fluorescence microscopy. (B) Flow cytometry analysis of cells expressing different levels of EGFR. Cells were incubated with 150 nM EGF-mCherry for 15 min before measurements. Folds of mean fluoresce difference with and without EGF-mCherry are presented for each cell line. (C) Internalization of EGF-mCherry was shown using A431 cells. Cells were incubated with 10 nM EGF-mCherry, fixed and counterstained with DAPI. (D) A soluble EGFR was immobilized on an ARG2 sensor for biolayer interferometry (BLI). Different concentrations of EGF-mCherry were measured and the data was fitted.

Biologic activity and specificity in receptor binding was verified for EGF-mCherry. Next, we established cellular assays that allow for characterization of binding inhibitors and blocking antibodies for the EGFR tumor target. A431 cells were incubated with elevating concentrations of EGF-mCherry at 37°C for 15 min and the mCherry signal was detected using a microplate reader. As already seen in previous experiments, incubation at this temperature and time interval comprises receptor binding, internalization and depletion. Therefore, besides the bound EGF-mCherry, the already internalized EGF-mCherry is also detected. Plotting the fluorescence intensity against the EGF-mCherry concentration shows a sigmoidal curve with signal or receptor saturation, respectively (Figure 10A).

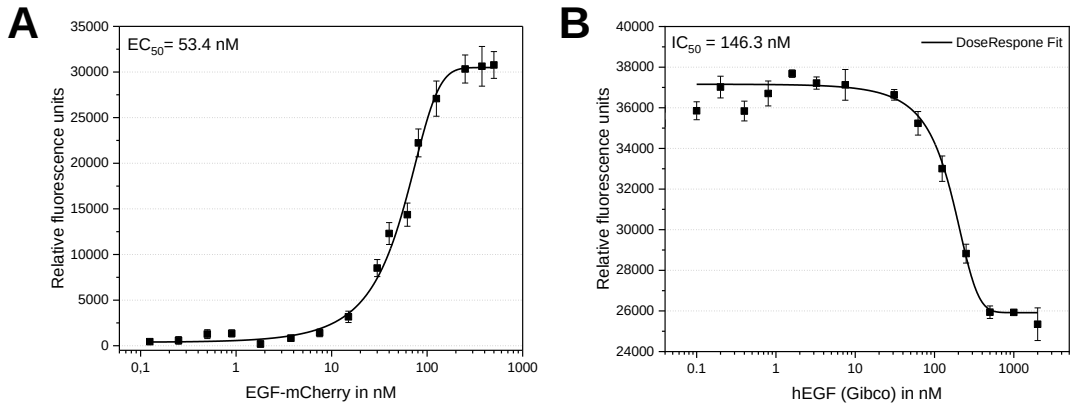


Figure 10: EGF-mCherry cell binding and competition assays. (A) A binding assay was performed on A431 cells that were incubated with increasing concentration of EGF-mCherry. (B) In the competition assay cells were simultaneously incubated with EGF-mCherry at 50 nM and hEGF in increasing concentrations. Relative fluorescence of cells was measured using a Tecan Spark instrument at 550/610nm. The data was fitted using a Dose-Response Fit in Origin2019.

The half-maximal effective concentration EC_{50} of 53.4 ± 4.2 nM was determined from a sigmoidal curve fitting. From this first binding experiment, the required concentration of EGF-mCherry can be derived for the following competition experiments. A concentration of 50 nM EGF-mCherry was kept constant while a non-fluorescent competitor, e.g. hEGF (Gibco) was added in increasing concentrations. Dose-response curve fitting of the fluorescence intensities allows for estimation of an inhibitory concentration IC_{50} for hEGF, which was determined to be 146.3 ± 36.2 nM (Figure 10B). The IC_{50} values would match the concentration of EGF-mCherry if both EGF variants were equally active. Here, the inhibitory concentration is three times higher, which might be explained by the observed formation of multimers for EGF-mCherry. In summary, these assays were easy and simple to implement in laboratories that are equipped with a standard microplate reader with an excitation in the range of 560 nm and emission above 580 nm and allow for characterization of EGFR-expressing cell lines and EGF-binding inhibitors.

5.2. Peptides binding the EGFR dimerization arm

Results of this project were summarized in a manuscript with the title ‘Designed EGFR-inhibiting peptides suitable for tumor-targeting of recombinant adeno-associated virus capsids’. The original manuscript is included in the appendix and a detailed summary of the work is presented in the following chapter.

Targeting of cancer cells overexpressing the EGFR was a main aim of this work. The target site for binding was selected previously with the EGFR dimerization interface. Since this interface is mainly involved in receptor dimerization, binding would not only allow for internalization into the target cell, but also in blocking of the receptor dimerization and activation. We focus on peptide

ligands that interact with the dimerization interface. These ligands were designed using a rational process relying on the crystal structure of the EGFR extracellular domain (PDB ID 1IVO). PepEDA was extracted from the dimerization arm structure of the EGFR and was used as a control as it was shown previously that it interacts with the EGFR (Figure 11).[85,86] Pep1osy and Pep1jhf were selected from a rational design approach that started with the extraction of 200 12-mer β-hairpins from the PDB. The extractions were further superimposed on to the native EDA hairpin (pepEDA). A rationally and translationally sampling into the binding pocket of the receptor was performed which resulted in 200,000 sequence/ position solutions. These solutions were filtered for satisfied hydrogen bonds and significant hydrophobic surface areas and afterwards screened for binding energies as predicted with Rosetta. A final selection of 30 solutions was visually inspected and four top solutions were chosen. These underwent additional high-resolution docking and sequence improvement processes. For final experiments the two top solutions pep1jhf and pep1osy were chosen which are presented in Figure 11.

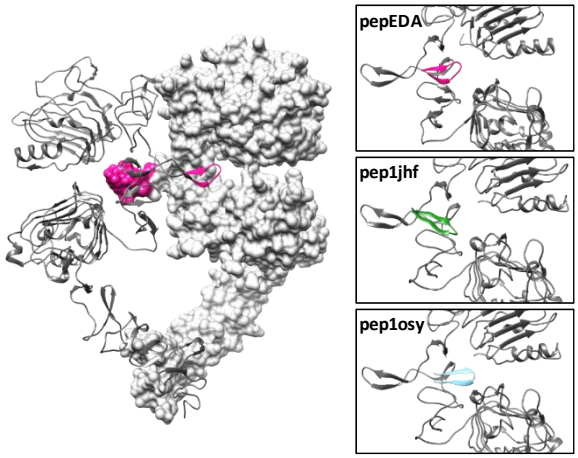


Figure 11: Crystal structure representation of the extracellular EGFR domain. Crystal structure of the dimerization complex of the human EGFR extracellular domains (PDB 3NJP). The EGFR dimerization arm (EDA) is highlighted in magenta in both domains. Rationally designed peptides binding the dimerization interface are highlighted in green (pep1jhf) and light blue (pep1osy). The rational design was performed by D.J. Mandell. Models were generated using UCSF Chimera.

Both top solutions and the control pepEDA were sought to be analyzed for EGFR binding but not only in the context of their individual affinity towards this receptor but also in terms of avidity which is achieved by incorporation of the peptides into the rAAV capsid.

5.2.1. Synthesis of cyclic EGFR binding peptides

The individual affinity of pep1jhf and pep1osy was to be tested for the synthetic cyclic peptides in direct comparison with EDA2. To study the rationally designed peptide ligands with literature described values, first a strategy taken from the literature was used to generate cyclic peptides. The peptide EDA2 was described in 2015 by Hanold et al. and represents the natural EGFR dimerization arm (EDA).[85] Cyclization of this peptide was achieved via a triazolyl bridge between the terminal amino acids. The formation of this bridge was tested using different terminal amino acids and a

combination of a *C*-terminal 4-azido-L-homoalanine (L-Aha) and an *N*-terminal L-progargylglycine (L-Pg) resulted in a triazole with a length closely related to the wild-type EDA.[85] This strategy was transferred to pep1jhf and pep1osy and peptides were synthesized on a Rink amide resin.

Table 2: Peptide sequences for generation of triazole-bridged EGFR-binding peptides.

Name	Sequence
pep1jhf	Pg-WAKSQGNKSEYQ-Aha
pep1osy	Pg-VVFEVNGRDLGW-Aha
EDA2	Pg-LYNPTTYQMD-Aha

First cyclization attempts were carried out on-resin as described in literature. Applying $CuSO_4$ (30 equiv)/ascorbic acid (5 equiv) in a solvent system of ^{t}BuOH:H_2O (1:2 (*v/v*), overnight) did not results in the observation of the final product. Instead, the starting material was not converted according to HPLC and IR. A second approach using Cu^I instead of Cu^{II} was applied with CuI (0.4 equiv)/ascorbic acid (0.8 equiv)/2,6-lutidine (1.6 equiv) in NMP/H_2O (4:1 (*v/v*), 96 h).[102] Successful click-chemistry between the terminal amino acids will not result in a shift in mass and thus, only a shift in retention time and IR spectroscopy can confirm a successful cyclization. The observed shifts in the retention time were very small and therefore IR measurements were performed. For a successful cyclization these should result in disappearance of the vibrations of the alkyne (2100 cm^{-1}). A disadvantage of this method, however, is that no reliable results could be obtained from crude, low-concentrated samples. Thus, IR measurements were conducted after final deprotection, cleavage from the resin and purification. Cyclization on-resin was not detected for peptides pep1jhf and pep1osy after both cyclization attempts (Figure 12). For the literature described cyclic pepEDA, the alkyne vibration was not detected and a shift in retention time was as expected observed in HPLC.

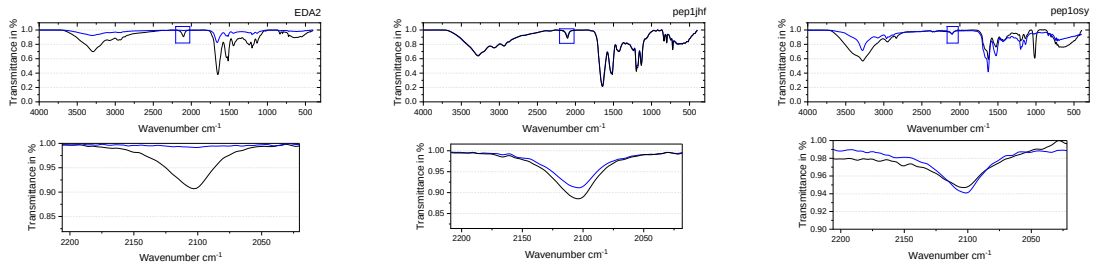


Figure 12: IR measurement of for linear (black) and potential cyclic (blue) pepEDA, pep1jhf and pep1osy. The characteristic vibration of the alkyne is seen for the linear as well as for the cyclic peptide. Measurements of a small peptide sample were conducted in methanol with a Bruker Alpha II Platinum ATR system. Data was acquired using OPUS software and analyzed in Origin2019.

Problems in on-resin cyclization can be related to different aspects. A previous publication showed that resins with long PEG linkers can contribute to the click reaction, because they allow more degrees of freedom.[85] Also, the peptide sequence was shown to have an influence on the outcome

of the click reaction. Since all peptides were already cleaved from the resin and purified we changed the strategy towards in-solution click chemistry. Three different test reaction were set up including $CuSO_4$ (4.4 equiv)/ascorbic acid (6.6 equiv) in $^tBuOH/H_2O$ (1:2 (*v/v*), overnight)[103], CuI (10 equiv)/DIPEA (25 equiv) in acetonitrile[103] and Cu(0) powder (15 equiv) in $^tBuOH/H_2O$ (1:2 (*v/v*), overnight). Again LC/MS and IR measurements were conducted but cyclization was not observed under these varying conditions. All attempts to carry out on-resin and in-solution intrachain Cu^I-catalyzed azide-alkyne 1,3-dipolar cycloaddition failed for pep1jhf and pep1osy. Literature suggested that the chemical nature of the linear precursor plays an important role for the propensity of the click-reaction.[103] Distance and chemical surrounding of the azido and alkynyl groups are important as well as the reaction conditions.[103] The length of the linear precursor (14 residues including L-Aha and L-Pg) was the same for pep1jhf and pep1osy but two residues longer compared to pepEDA. This difference in distance probably has the greater influence on the reaction, because the sequences of the peptides are substantially different and thus should influence the reaction not in the same way.

As this method was running into a dead-end, alternative cyclization strategies had to be considered. In 2016, Toyama et al. analyzed a cyclic peptide, close to pepEDA, and showed its biological function.[86] The decapeptide CYNPTTYQMC was cyclized via a disulfide bridge, that links the two terminal cysteines. This peptide was chosen as a new model peptide for comparison with pep1jhf and pep1osy. Thus, the literature described pepEDA, as well as pep1jhf and pep1osy were synthesized each with terminal cysteines in a linear, a cyclic and a 5(6)-carboxyfluorescein-labelled cyclic variant. Disulfide-cyclized peptides were synthesized and analyzed in cooperation with Isabell Kemker (Prof. Dr. N. Sewald).

5.2.2. Pep1osy reveals β-hairpin character

Rational peptide design was performed to result in β-hairpin-structured peptides that are able to bind subunit II of the EGFR. We obtained synthetic disulfide-bridged peptides that needed to be analyzed for their structural properties. The reference peptide pepEDA in the triazolyl-bridged and in the disulfide-bridged form have previously been analyzed for secondary structural characteristics in circular dichroism (CD) experiments, proving formation of secondary structural elements.[85,86] Estimating the secondary structural characteristics for the rational designed peptides was carried out using circular dichroism experiments in comparison to pepEDA. All peptides were solubilized in 10% trifluoroethanol (TFE) in water and measured subsequently in comparison to their linear progenitor. Results are shown in Figure 13 for the cyclic peptides in direct comparison to the linear progenitor. Only pep1osy shows a significant β-hairpin character with a maximum at ~195 nm and a minimum at ~210 nm in comparison to the linear precursor that shows an unfolded structure with a minimum at approximately 200 nm. PepEDA in comparison indicates some low levels of structuring. In contrast, the linear peptide exhibits a more flattened profile typical of a flexible disordered structure with a minimum at approximately 195 nm. In the context of the disulfide-bridged context

no secondary structure formation was observed for pep1jhf as the anticipated shift towards higher wavelengths was not observed. Here, the secondary structure of a β-hairpin was not confirmed under several tested conditions (different buffer compositions). The presence of the cyclic and linear form was verified using LC/MS directly after CD measurements, thus we think that pep1jhf remains a random coil conformation under these conditions.

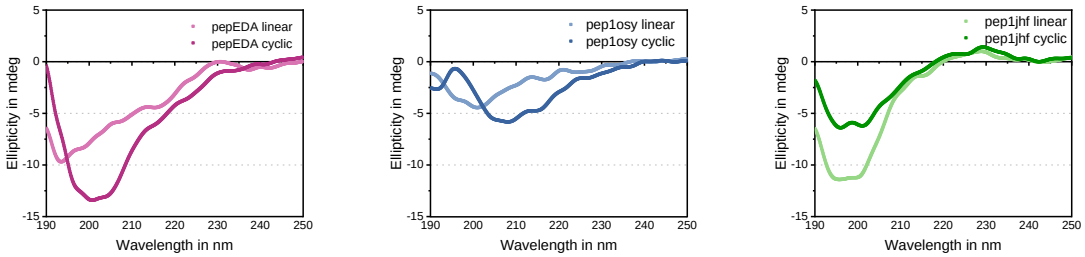


Figure 13: Circular dichroism (CD) spectra of linear and cyclic peptides. CD spectra for all linear and cyclic peptides were recorded using a Jasco-810 from 190 to 250 nm wavelength at 25 °C. Peptides were dissolved in 10% trifluoroethanol (TFE)/water (v/v) at 100 µM concentrations.

Nonetheless, all further experiments were performed using all three cyclic peptides. Even if no secondary structure is observed, the cyclic peptide might have the ability to bind the extracellular domain of EGFR.

5.2.3. Pep1osy shows inhibitory properties in wound-healing assays

We hypothesize, that upon binding of the peptide to subunit II dimerization of the EGFR will not occur because the dimerization interface is blocked. Testing the influence of the peptides was assayed in wound healing assays. During these experiments closure of a wound is monitored and cell migration is analyzed. Several signaling pathways activated by the EGFR are involved in cell migration and adhesion and thus, peptides should affect receptor dimerization. These effects should prohibit the migration induced by EGF and result in cell migration rates comparable to those of the buffer controls. Assays were performed on A431 cells that overexpress the EGF receptor and results are presented in Figure 14A, B. In the buffer only control experiment cell migration rates are low with about 10 µm per hour. In contrast, the experiment in presence of 1 nM of the natural EGFR ligand human EGF (hEGF) yielded elevated migration rates up to three times, proving the strong impact of hEGF. Peptides were applied to the cells in 5 µM concentration and showed migration rates comparable to those of the negative control. This was expected as cells were serum-starved prior to incubation and thus, no ligand should be available to induce dimerization. In the second setup, peptides (5 µM) were simultaneously incubated with 1 nM hEGF. Here, hEGF binding would allow for dimerization. Blocking of the receptor due to the peptide would inhibit dimerization and thus, signaling inside the cell will not activate pathways involved in cell migration. pepEDA and pep1jhf show a retained high cell migration rate in presence of hEGF. For pep1jhf this result could be expected as the secondary structure was not confirmed in CD experiments. For pepEDA this results is surprising, as in previous publications the downregulation of the EGFR was shown in

Western blot analyses.[86,88] Only for pep1osy we clearly saw a decrease in cell migration velocity in the presence of hEGF which was detected in the range of the buffer control. Here, we hypothesize that binding of the peptide corresponds to an inhibition of dimerization and further to a downregulation of cell signaling pathways.

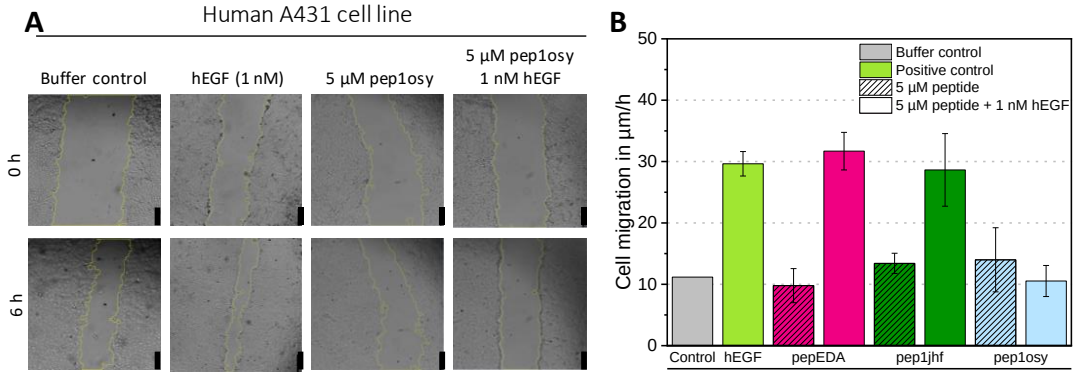


Figure 14: Wound healing assay on A431 cells. Cells were seeded in 24-well plates, cultivated until they reached confluence and serum-starved prior to creation of wounds. After extensive washing cells were incubated with peptides. Images were taken hourly using a Leica DMI6000 and cell-free areas and cell migration were determined using the ImageJ software. Scale bars represent 250 µm. (A) Images of wounds in the negative buffer control, the positive 1 nM hEGF sample, 5 µM pep1osy and 5 µM pep1osy in combination with 1 nM hEGF. (B) From plotting of the wound area against time, cell migration rates were determined by applying a linear regression in Origin2019. Cell migration rates were determined for biological replicates.

5.2.4. Cyclic peptides bind and internalize into EGFR overexpressing cells

Besides the cyclic peptide variants already tested, fluorescein-derivatives have been synthesized of pepEDA, pep1jhf and pep1osy. These fluorescent-labelled peptides were used to assay cell binding properties on cells expressing different levels of EGFR. Experiments were primarily conducted using EGFR-overexpressing A431 cells and low-level expressing MCF7 cells. In the past, both cells lines have been extensively studied regarding their EGF receptor state.[101,104]

Combining flow cytometry and confocal laser scanning microscopy on live cells allows for assessment of peptide import. Confocal microscopy allows subcellular localization of peptides but is not quantitative, while flow cytometry is quantitative but does not provide spatial information. The first experiment performed with both cell lines was a flow cytometry-based assay. Cells were incubated with peptides at a 20 µM concentration for 15 min at 37 °C and after washing steps a subsequent flow cytometry analysis was performed. In this assay a fluorescence signal can be related to peptide bound to the cell surface as well as internalization into the cell. Furthermore, it cannot distinguish between receptor-mediated or random internalization. Data in Figure 15 show the normalized mean fluorescence intensity, demonstrating the increase in fluorescence for all peptides on both cell lines. Interestingly, this increase is also visible for the low-level EGFR expressing MCF7 cell line. But in comparison the fluorescence signal detected on A431 cell is elevated for all three peptides indicating that here high amounts bind and internalize into cells.

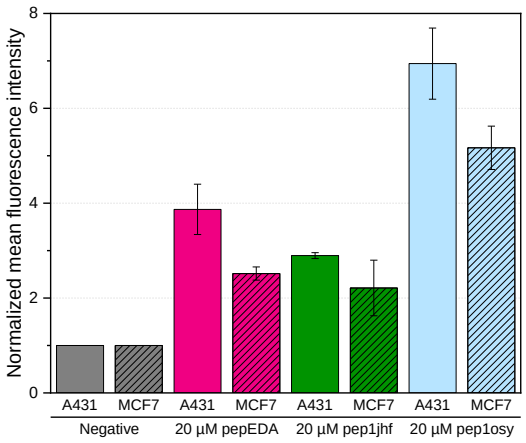


Figure 15: Flow cytometry analysis of MCF7 and A431 cells with pepEDA, pep1jhf and pep1osy. Cells were seeded the day before analysis on 24-well plates at a density of 1×10^5. Peptide were diluted to 20 µM concentration in medium and incubated for 15 min at 37 °C before detachment and analysis using a FACScalibur. 10,000 events were counted and data analysis was performed using FlowJo and Origin2019.

Further analysis of cell internalization was performed using confocal laser scanning microscopy. Here, peptide internalization was tracked after addition of 5 µM fluorescein-labelled peptide for 10 min and prior staining of acidic cell compartments using Lysotracker DND-99. The EGF receptor is known to internalize via the clathrin-mediated pathway in endosomes and ends up in lysosomes where degradation occurs.[105] If the carboxyfluorescein-labelled peptide is internalized into cells via endocytosis this can be tracked via colocalization of the fluorescence from both, the peptide and the lysotracker. Figure 16 shows fluorescence images for all cyclic peptides on MCF7 and A431 cells. Colocalization was analyzed using a self-designed macro in ImageJ and is presented in white spots in the fluorescence image overlays. For better visualization white arrows indicate areas where colocalization was detected. Similar low levels of colocalization of pep1jhf and pepEDA were identified which correspond to the flow cytometry data and the low β-hairpin character observed for both peptides. Here a more diffuse, no lysosome-related mechanism leads to internalization into the cell. For pep1osy a more specific internalization was observed also represented the flow cytometry results which can be attributed to the strong β-hairpin character observed in CD spectra.

Summarizing these results, we believe that internalization of pep1osy utilizes a specific EGFR-related cell entry mechanism. Analysis performed for pep1jhf and pepEDA verified all previous results and shows that the β-hairpin secondary structure is essential for true receptor-mediated internalization.

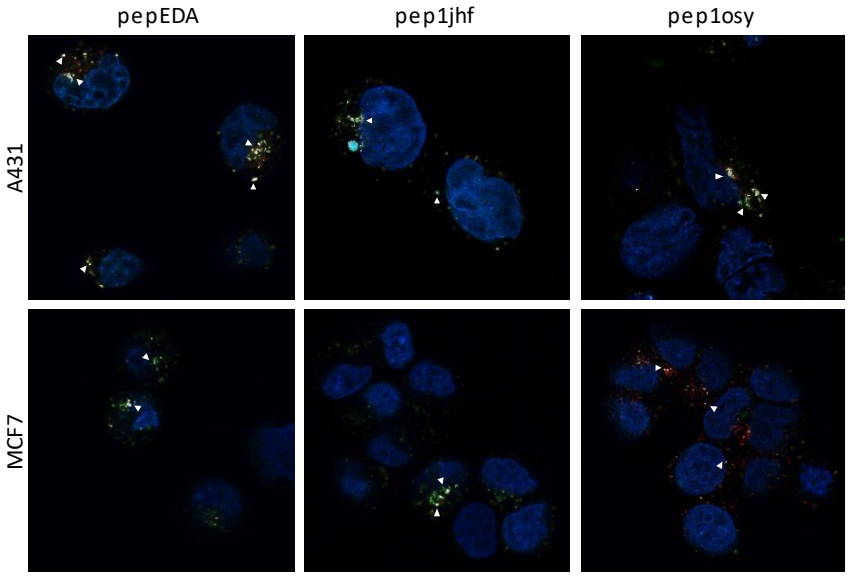


Figure 16: Live cell imaging of cyclic peptide variants on A431 and MCF7 cells. Fluorescence microscopy was performed using an inverted laser scanning Zeiss LSM780 microscope. Cells were seeded on LabTek slides the day before the experiment. Nuclei were counterstained using NucBlue and lysosomes were stained using Lysotracker DND-99 before incubation with 5 µM fluorescein-labelled peptide for 10 min at 37 °C. Images were analyzed using Zeiss Zen 2011 and ImageJ software. Colocalization of peptide and Lysotracker DND-99 was analyzed in ImageJ and all spots of colocalization are marked in white and highlighted with white arrows.

5.2.5. Pep1osy shows an affinity towards the soluble EGFR

Binding affinities of the peptide were now analyzed in fluorescence polarization assays using the soluble EGFR variant described in the previous chapters. The carboxyfluorescein-labelled peptides were each analyzed for their affinity towards the receptor without addition of the natural ligand EGF. Excitation of a fluorophore with linear polarized light results in emission of linear polarized light in dependence of its fluorescence half-life and rotation speed. At constant temperature and viscosity, fluorescence polarization depends on the volume of the fluorophore-labelled small molecule. If a fluorophore-labelled ligand is free in solution, its volume is small and its mobility is high, depolarization of the emitted light is strong. If this ligand binds to the target protein, its volume and mobility decreases due to limited rotation, whereby the emitted light remains more polarized during the fluorescence lifetime.[106,107] After incubation of peptide:EGFR solutions, fluorescence polarization was measured using a Tecan Spark instrument. Polarization was converted to anisotropy and plotted against the EGFR concentration (Figure 17A,B). The data were analyzed according to a 1:1 binding model equation described in literature.[108] For pep1jhf and pepEDA no increase in fluorescence anisotropy was observed in the accessible concentration range for EGFR. Only pep1osy showed an increase which was fitted with the described binding model resulting in an apparent dissociation constant of ~70 µM. Data points are missing a plateau phase at increasing sEGFR concentrations, which needed to be assumed for the fitting model (black data points). A value of 1 M EGFR was assumed as EGFR concentration at a maximum anisotropy of 200 mA, since in previous experiments with proteins and peptides of the same molecular weight a difference of about 150 mA between minimal and maximal anisotropy was observed. A plateau could not be

achieved, because the material input is not related to the added value of the data. Compared to other EGFR binders this value is considerably low showing that the affinity towards the dimerization arm limited. The pattern of binding is more complex at it seems initially. The peptide binding site in domain II is only accessible in the extended conformation of the receptor. The tethered structure shows an intramolecular interaction between subunit II and subunit IV which is potentially preferred over the dimerization, as domain IV has a higher local concentration (Figure 17C). Preformation of a receptor dimer in absence of a ligand has not been described.[109] In presence of a ligand, dimerization occurs with a K_D~1 µM.[110] Ligand binding allows a transition into the extended conformation and thus, also a clear shift of the equilibrium in the direction of dimerization. Performance of this experiment with the addition of the natural ligand EGF would not have simplified this experiment, as dimerization would again be preferred and in direct competition with the binding of the EGFR-specific peptide. Hence, the direct binding model by Roehrl et al. does not describe this relationship in full detail and we hypothesize that the real dissociation constant for the peptide is even lower.

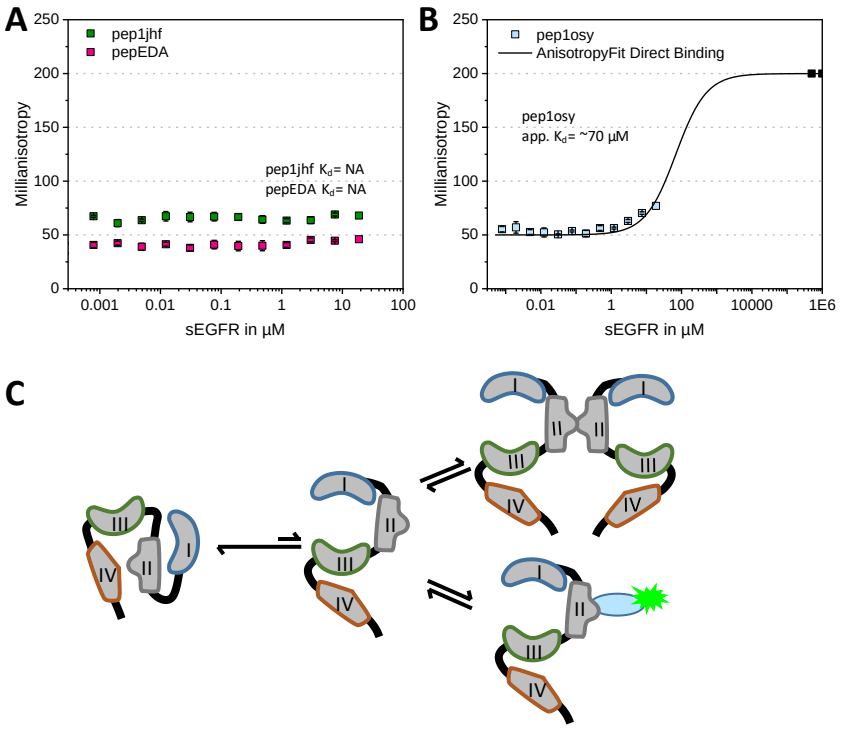


Figure 17: Fluorescence polarization assay for determination of the dissociation constant for binding of the fluorescent-labelled peptides to the soluble EGFR. (A) Normalized anisotropy of pep1jhf and pepEDA was plotted against EGFR concentration. Fitting was not successful as no change in anisotropy was observed. (B) Anisotropy data for pep1osy shows an increase which was fitted according to equations described in literature by Roehrl et al..[108] Assuming an anisotropy maximum of 200 mA at 1 M sEGFR (black points), a dissociation constant of 68.6 µM was determined. Errors were calculated for three independent triplicates using Gaussian error propagation. (C) Complex binding pattern during fluorescence polarization assays. The tethered form of the EGFR is not able to bind the fluorescent-labelled peptide. Preformation of a receptor dimer competes with binding of the fluorescent-labelled peptide.

Results of this assay summarize the previously observed findings. The natural derived peptide pepEDA showed no detectable binding to the receptor, which might be attributed to a low binding affinity. Likewise, no binding affinity was observed for the rational designed peptide pep1jhf. Here

the peptide might, as also seen in other assays, not be present in the correct β-hairpin conformation and binding to the receptor is inhibited. As already observed in the previous assays, pep1osy showed an affinity towards EGFR. Nonetheless, the K_D is considerably low and random internalization was observed. In order to avoid these problems, a system was found with the AAVs in which a high number of peptides increases the avidity and the larger size also prevents random internalization.

5.2.6. Integration of peptide ligands into rAAV2

Analysis of the synthetic peptide showed interesting features of the rationally designed EGFR-binding peptides in comparison with the already characterized variant pepEDA. Our aim was to incorporate all peptides into the capsid structure of an adeno-associated virus. Here, the 587-loop region of the VP proteins was selected for the site of integration resulting in 60 modified viral proteins composing a capsid Figure 18A. Increasing the number of available ligands with low affinity increases the avidity of the overall molecule. We aimed at incorporation of the peptides into the capsid to selectively target cancer cells expressing high levels of EGFR and deliver a gene of interest (GOI) into the cell for therapy. In the context of the synthetic peptide, we evaluated the ability of internalization which is an important feature for internalization of the whole capsid into the tumor cells.

As shown in Figure 18A incorporation of peptide ligands at residue position 587 of the VP protein occurred on a genetic level. Here, amino acid sequences were introduced without the addition of terminal cysteine residues not required for interaction with EGFR. As starting point, a plasmid coding for the capsid proteins of AAV serotype 2 was used (pZMB0216) that enables the insertion into the 587 loop region via unique restriction site BamHI and PvuII. Hybridized oligonucleotides were used to incorporate the right DNA sequence after restriction digest of the starting plasmid. To prevent resulting rAAV2 from transducing cells via their natural tropism two amino acid mutations were further introduced (R585A, R588A). Both amino acid residues are known to mediate the contact of the virus capsid towards HSPG and mutation to alanine's was shown to dramatically reduce the ability to successfully transduce cells.[111] The position of amino acids mutations is highlighted in Figure 18A in context of the crystal structure (left) and amino acid sequence (right). Plasmids with integrated sequences were verified for their correctness via Sanger sequencing before application in rAAV2 production. Ten 100 mm plates were used in a standard rAAV production in adherent HEK293 cells. The final titer of a sample concentrated at about 500 µl was determined by qPCR by direct comparison to a standard curve. As seen in Figure 18B final titers of all produced rAAV variants are equivalently high showing that introduction of peptides into the 587 position does not interfere with capsid assembly. The final preparations were also assayed for their correct VP protein ratios in a Western blot (Figure 18C) proving that the normal molar ratio of 1:1:10 for VP1,VP2 and VP3 is maintained.

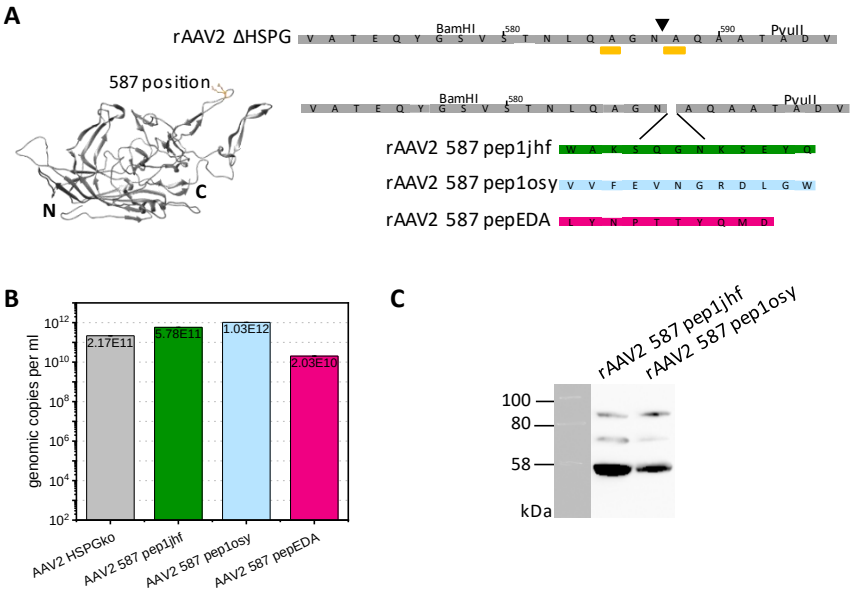


Figure 18: Incorporation of EGFR-binding peptides into rAAV2. (A) Representation of the VP protein with yellow highlighted integration site at amino acid residues 587 (PDB 1LP3). Sequences of variable region VIII and of all peptides are presented and the integration site shown. Yellow bars indicate R585A and R588A mutations leading to a HSPG knock out. (B) Genomic titers of rAAV2 peptide variants determined via qPCR. Yields were obtained from 10× 100 mm dishes after purification using iodixanol ultracentrifugation. (C) Western blot analysis of rAAV2 587 pep1jhf and rAAV2 587 pep1osy. VP proteins were detected using the anti-VP antibody B1, which is detected in a chemiluminescence reaction upon binding of the secondary HRP-coupled antibody.

All produced rAAV variants carry the expression cassette of the fluorescent protein mVenus with a CMV promoter in between the ITRs. If the peptides mediate transduction, lysosomal escape and transgene delivery to the nucleus of EGFR-overexpressing cells, expression of the fluorescent protein mVenus takes place. Thus, analysis of successful transduction was observed via flow cytometry after cells were incubated with a multiplicity of infection (MOI) of 50,000 for a duration of 96 h. Here, four rAAV2 variants were analyzed for their transduction ability on A431 and MCF7 cells in comparison to a buffer control (negative). The transduction ability of rAAV2 HSPGko is in the range of the negative control indicating that two arginine mutations to alanine reduce the transduction ability dramatically. Nonetheless, a slight difference can be observed between the two cell lines. The wild-type rAAV2 is known to transduce several cell lines to very high extents.[112] Control experiments with wild-type rAAV serotype 2 are included in chapter 5.3.1. For the pepEDA insertion in rAAV2, transduction efficiency is located in the range of the negative control showing that retargeting of the EGFR is not reached by integration of the natural sequence of the EGFR dimerization arm. As shown in chapter 5.2.4, this peptide is able to internalize into both cell lines, but transduction of cells requires also a successful lysosomal escape as well as delivery of the transgene into the nucleus for gene expression. If one of these steps is prohibited, mVenus will not be expressed, even if internalization occurred. Integration of rational designed pep1osy and pep1jhf resulted in about 30% of mVenus positive cells proving a successful transduction of A431 while for MCF7 cells the percentage of fluorescent cells is in the range if the negative control. By comparing the data of the synthetic peptide with the data from rAAV insertion this was not expected

as for pep1jhf no clear secondary structure could be formed. A possible explanation can be given by the β-hairpin structure that is already present in the loop region 587 of the VP protein. During the design of the cloning, care was taken to ensure that this structure is retained even after peptide introduction. The rigid fixation may also result in pep1jhf showing this secondary structural element. In context of rAAV, this peptide was shown to be functional. In general, increasing the avidity of the peptide, by presentation on the capsid surface, and the size of the whole capsid decreases the amount of unspecific internalization in comparison to the synthetic peptide and thus, transduction should only occur via receptor-mediated endocytosis of the EGFR.

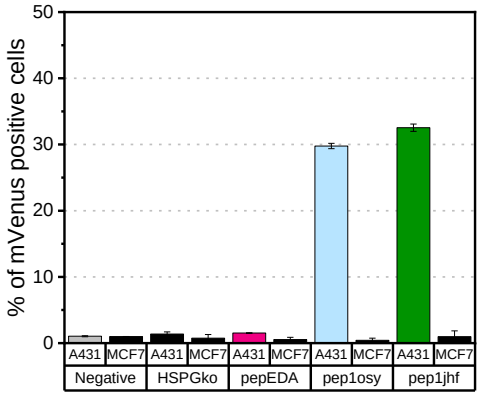


Figure 19: Results of transduction experiments using rAAV 587 peptide variants on A431 and MCF7 cells. A431 and MCF7 were seeded, rAAV samples at a MOI of 50,000 were applied and incubated over 96 h at 37 °C. Successful transduction is verified by expression of the delivered gene of interest mVenus which is under the control of a CMV promoter. Fluorescence was measured using flow cytometry at a FACScalibur by counting 10,000 events for biological duplicates. Data analysis was performed using FlowJo software and Origin2019.

5.2.7. Assaying AAVR dependency of rAAV2 variants

We were further interested on the dependency on the AAV receptor (AAVR or KIAA0319L). This receptor was described in literature to be necessary for AAV2 transduction.[40] Introduction of retargeting motifs and usage of a different internalization pathway might influence dependency on the AAVR. To investigate the influence on transduction, a stable cell line was generated using CRISPR/Cas9 to induce a AAVR knock out. The CRISPR target sequences for exon 2 was chosen from literature but was until now not tested for this cell line (Figure 20A).[40] The target sequence was cloned into pSpCas9(BB)-2A-GFP (PX458) from the Zhang laboratory[113], transfected and cells were subsequently sorted for EGFP expression using a FACS S3e system (BioRad). Monoclonal colonies were cultivated and genomic DNA isolated before subcloning and subsequent analysis via Sanger sequencing. A431 cells are diploid for chromosome 1p34-36 and a total knock out both copies of the AAVR needed to be affected. For one A431 clone a 20 bp deletion (del20) leading to a frame-shift and complete loss of the protein was observed. For the second copy a 36 bp deletion (del36) in the exon-intron junction was detected. As this represents a multiple of three, a total loss cannot be excluded, but due to the position on the intron-exon junction it was expected, that at least

a loss of domains occurs, resulting in a stable A431 cell line with extremely reduced AAVR expression. Furthermore, the first transmembrane domain is located in this area, which is severely affected by the large del36 mutation.

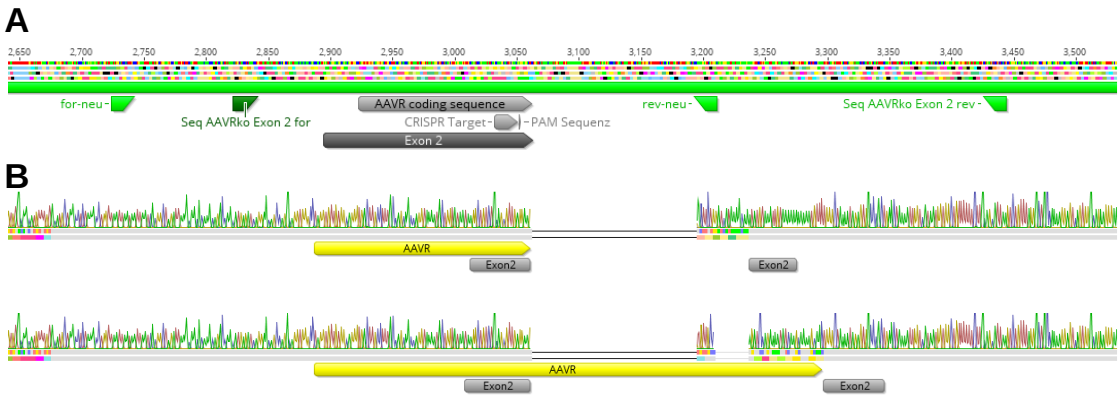


Figure 20: Generation of a A431 AAVRko cell line using CRISPR/Cas9. (A) Genomic sequence of Exon2 of chromosome 1p34-36. Showing the CRISPR target site (sgRNA sequence) and the protospacer adjacent motif (PAM) sequence for induction of the AAVR knock out. Positions of sequencing primers are given as green bars. (B) Results of Sanger sequencing analysis after extraction and sub-cloning of A431 genomic DNA showing that two events occurred either del20 (top) or del36 (bottom).

Engineering of the A431 cell line might influence the EGF receptor level on the cells surface. Thus, EGF-mCherry was used to determine the EGFR expression level. Cells were incubated with 100 nM EGF-mCherry for 15 min before analysis using a FACScalibur system. Analysis revealed that the mean fluorescence intensity is located in the same range for both cell lines, proving that the EGFR is still presented in the surface. Further transduction experiments were conducted using a previously described rAAV variant.[60] Here, a DARPinE01 was introduced in the rAAV2 capsid. DARPins are synthetically designed proteins that can be selected to bind different targets. Incorporation of these 14-17 kDa proteins into the rAAV capsid was shown to be tolerated by *N*-terminal fusion to the VP2 protein.[59,60] Resulting viral particles present the DARPinE01 in all VP2 proteins, meaning about five incorporations within the whole capsid structure. This variant was used as a positive control as various working groups have demonstrated its ability to bind and internalize into cells. As represented in the heat map of Figure 21, the AAV2-DARPin variant shows selective transduction of A431 cells in comparison to wild-type AAV2. In this assay also the transduction ability on the novel created A431-AAVRko cell line was evaluated proving that in comparison to the parental A431 cell line the transduction ability is dramatically reduced. Nonetheless, a complete reduction was to be expected, which was not observed. This might be explained by the del36 mutation, which does not lead to a full eradication of the AAVR. We also tested AAV2-pep1osy's transduction ability on the A431-AAVRko cell line in comparison to the parental A431 and the MCF7 cell line. Previous results could be verified showing a specific transduction of A431 cells but no fluorescent signal on MCF7 cells. For the A431-AAVRko cell lines a reduction in transduction ability was observed proving that besides the change in internalization pathway the AAVR protein still has an important role in the transduction process. It remains unclear at which point this protein takes part

in the transduction process, either as a cell surface receptor or as part of the intracellular processing.[114] However, the impact even on strongly modified rAAV variants with changed tropism is clearly observed in this experiment.

AAV2				
MOI 50,000	AAV2	AAV2 HSPGko	AAV2DARPinE01	AAV2 pep1osy
A431	52	1	86	58
A431 AAVRko	20	1	53	11
MCF7	53	1	8	0

Figure 21: Heat map of transduction results. Cell were seeded directly before transduction with rAAV variants at a MOI of 50,000. Transduction was analyzed for expression of the delivered transgene mVenus by flow cytometry using a FACScalibur. 10,000 events were counted and analyzed using FlowJo and Origin2019.

5.2.8. Testing serotype dependency of peptides

Next, we were interested if this retargeting mechanism is transferable to other AAV serotypes. A plasmid system was developed allowing for the introduction of peptide motifs into loop structure of AAV6 and AAV9 (detailed description can be found in chapter 5.3.2). For both serotypes, insertion took places at amino acid residue position 588 in correspondence to crystal structure analysis and comparison to AAV2 (Figure 22). Introduction of hybridized oligonucleotides was enabled by the restriction enzymes MscI and BamHI. Both serotypes have been described to tolerate modifications in this variable region. For AAV9 insertion at residue position 589 was described by Michelfelder et al.[61] We chose incorporation at 588 as we wanted to elongate this variable region under maintenance of the β-hairpin structure (Figure 22A). In comparison for AAV6 only an insertion at residue 585 was described for insertion of an RGD peptide.[62] Again, we chose incorporation at the 588 site to obtain the structural properties of the peptide and the loop region (Figure 22B).

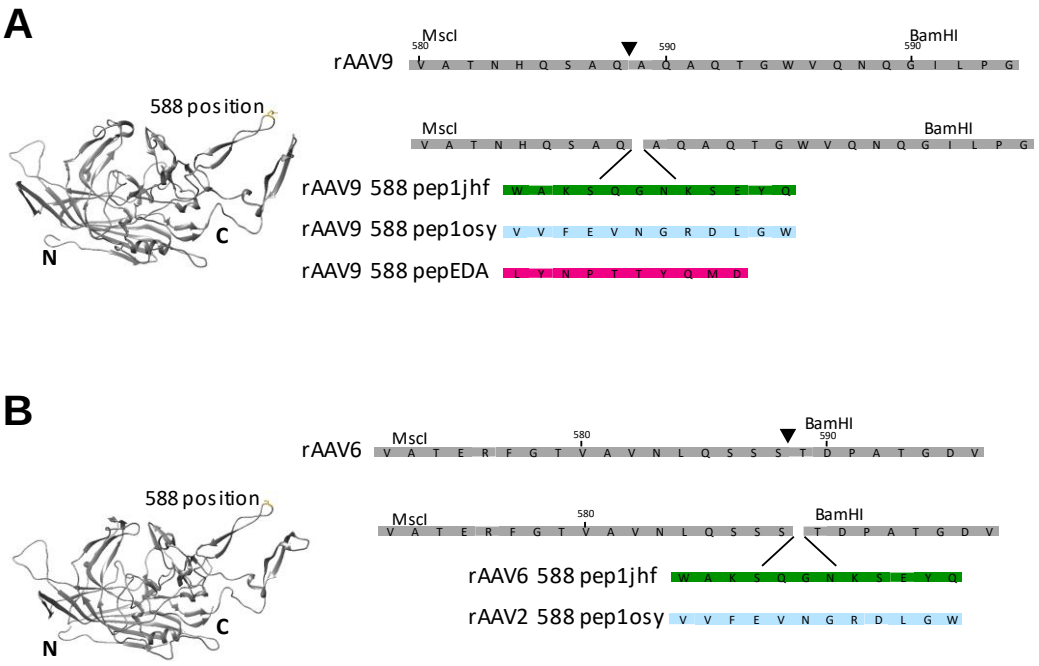


Figure 22: Incorporation of EGFR-binding peptides into rAAV9 and rAAV6. (A) Representation of the AAV9 VP protein with yellow highlighted integration site at amino acid residues 588 (PDB 3UX1). Sequences of the variable region VIII and of all peptides are presented and the integration site shown. (B) Representation of the AAV6 VP protein with yellow highlighted integration site at amino acid residues 588 (PDB 3OAH). Sequences of the variable region VIII and of all peptides are presented and the integration site shown.

Viral particles of all constructs were produced with satisfying yields and used for transduction of A431 and MCF7 cells at the same MOI as used for the previous rAAV2 experiments. Viral AAV9 particles with a wild-type capsid show, as expected, low transduction efficiencies in both cell lines (Figure 23).[112] For AAV9 pep1jhf a comparable efficiency was detected on MCF7 cells, while on A431 cells higher amounts of fluorescent cells have been detected. Here, an increase in specificity was observed compared to wild-type capsids. For rAAV9 pepEDA comparable low transduction results were obtained in either cell line. This was to be expected as insertion of this peptide was not introducing an EGFR-tropism in the AAV2 context as well. The most surprising results were obtained for rAAV9 pep1osy. A dramatic increase in transduction efficiency was observed in context of serotype 9. Even higher values were measured for MCF7 in comparison to A431 cells. In this setting the transduction properties improve surprisingly. We hypothesize, that properties of the incorporated peptide change within their surrounding structural context, leading to overall higher transduction abilities. This was also observed in the context of AAV6, which was chosen because of its known secondary receptor EGFR. The first idea was, that incorporation of EGFR-targeting peptide might increase the specificity of transduction. For AAV6 with a wild-type capsid nearly 100% transduction efficiency was observed on A431 and MCF7 cells. Incorporation of a hexahistidine-tag at the same position reduced the amount of fluorescent cells to the same extent. Incubating cells with rAAV6 pep1jhf leads to an even higher reduction in transduction ability compared to rAAV 587His. Comparable to the results for AAV9, this was not observed and might be related

to the structural properties of the capsid surrounding. In this context rAAV6 pep1osy also led to a reduction of transduction ability.

AAV6				
MOI 50,000	AAV6	AAV6 588 His	AAV6 pep 1jhf	AAV6 pep1osy
A431	99	16	2	9
MCF7	95	16	2	11

AAV9				
MOI 50,000	AAV9	AAV9 pep1jhf	AAV9 pepEDA	AAV9 pep1osy
A431	8	18	4	43
MCF7	7	8	4	60

Figure 23: Transduction analysis of rAAV9 and rAAV6 variants. A431 and MCF7 cells were incubated with a MOI of 50,000 and incubated for 96 h before flow cytometry analysis. 10,000 cells were analyzed for their mVenus fluorescence signal using a FACScalibur. Data was analyzed using FlowJo and heat maps were generated in Excel.

Overall these results suggest, peptides are not easily transferable from one serotype to another. Structural properties of the capsid play a major role in the definition of a tropism, which is not easily modified. Comparing these results, it needs to be outlined, that for rAAV2 the natural HSPG tropism was knocked out by introduction of two point mutations. Mutations leading to a dramatic loss of receptor affinity are also described for AAV9 and AAV6 and need to be implemented into this system to study the effects on re-targeting strategies by peptide insertion.

5.2.9. Wild-type rAAVs prove functionality in egg xenografts

Gaining deeper inside into peptide-modified rAAV vector variants, a more sophisticated approach besides cell culture is required. Chick chorioallantoic membrane (CAM) assays have been used in the past to study tumor growth and vascularization.[115–117] The CAM surrounds the fertilized chicken embryo and is composed of a multilayer epithelium.[118] After opening of the egg-shell, human cancer cells are transplanted on the CAM (Figure 24, Day 4 and Day 8). After four days incubation, a solid tumor is formed, which has a highly vascularized structure, multiple cell types and an extracellular matrix (see Figure 24, Day 12).[116] This rapid growth of a tumor with properties close to human tumors is a great advantage in comparison to mouse models, where time expenditure is a lot higher. Differences in time and effort are also seen in the cost of an experiment, which is considerably lower for the CAM model. We used the *in ovo* model to study the effectiveness of rAAV delivery to tumor tissue in collaboration with Dr. Lea Krutzke (University Hospital Ulm). We chose transplantation of A431 cells in this assay, as this epidermoid carcinoma cell line is known for a high EGFR expression level and previous assays show that rAAV transduction via the introduced peptide motifs is possible. The rAAV preparations were injected intravenously at 1×10^{11} viral vectors per egg on Day 12 and incubated for two more days. At Day 14 the tumor and the organs were isolated for analysis.

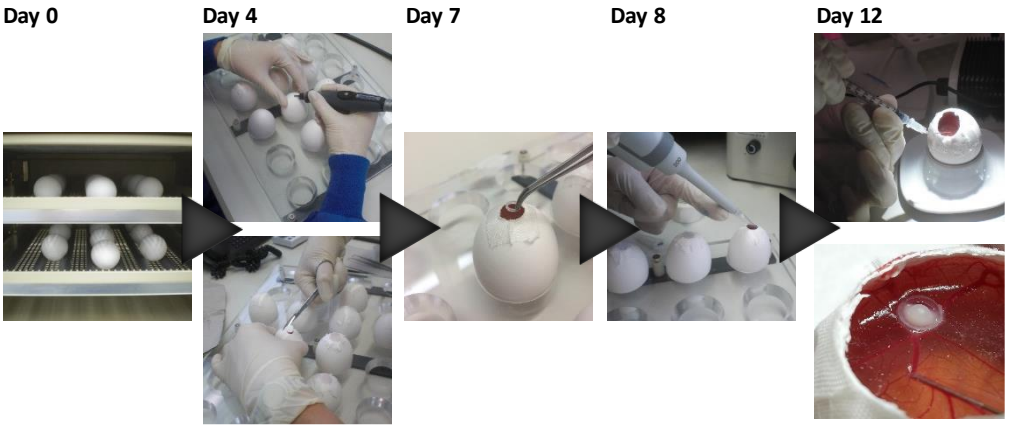


Figure 24: Overview on the in ovo chick chorioallantoic membrane assay procedure. The assay starts at day 0 with the arrival of the fertilized eggs. On day 4 the eggs are opened before on day 7 a silicon ring is placed on the CAM to provide a boundary for the tumor cells. The next day a suspension of tumor cells is added inside the silicon ring. A solid tumor had been formed until day 12 before injection of 1×10^{11} viral vector intravenously. On day 14 propofol is injected and organs of the embryo are isolated for further analysis.

Systemically delivered rAAV vectors that were stable in the blood circulation and were able to escape the vascular system have the chance to transduce tissues. *In ovo* CAM assays have, to our knowledge, not yet been described for AAVs. Thus, we were the first to perform such experiments and we wanted to see if this tumor model is comparable to existing platform technologies, e.g. tumor mouse model. Viral vectors were produced in adherent HEK293 cells and purified via iodixanol-gradient ultracentrifugation for rAAV2 and rAAV9 with a wild-type capsid. Vectors were injected systemically and both preparations showed no adverse effects in the chicken embryos after injection. Both rAAVs are packaging the gene for the fluorescent reporter protein mVenus which is under the control of a CMV promoter. Detection of fluorescence as well as determination of genomic copies from the extracted organ give information about the distribution of viral vectors in the tissues. Fluorescence intensity determination of the tissues was not possible due to unsuitable equipment. We only had the opportunity to characterize the genomic copy distribution in isolated organs. Here, DNA was isolated from the collected organ material and subsequently used as template in qPCR experiments. In qPCR experiments 20 ng isolated DNA was used per reaction with a primer pair specifically binding the mVenus gene. The house-keeping gene β-actin was used for normalization and results are presented in box-whisker diagrams in Figure 25. Eggs with injected PBS buffer serve as a negative control and are shown in comparison to rAAV2 and rAAV9. In general, we see that transduction efficiency of rAAV9 is elevated compared to rAAV2. This is in good agreement with data obtained in mouse experiments[119], but stands in contrast to experiments performed in cell culture experiments[112]. These experiments showed, that in cell culture transduction efficiency of AAV2 outweighs that of AAV9. Due to fast blood clearance and poor transcytosis AAV2s transduction efficiency remains low in *in vivo* experiments.[120] In comparison AAV9 shows slow blood clearance and high transcytosis resulting in a longer persistence in the embryo. By gaining a deeper insight into the distribution of viral genomes, the expected tropism for certain organs, which corresponds to that in mice, becomes clear. AAV2 was reported to transduce a broad

range of tissue types but shows elevated transduction of kidneys. This was also observed in the CAM assay, where the kidney was the predominant target of the vector (Figure 25E). For AAV9 transduction a similar but more efficient profile was observed but here especially the liver and the brain were described to show improved transduction.[121] Targeting of the brain is possible because it was described that AAV9 is able to cross the blood-brain barrier via transcytosis, while AAV2 is only able to transduce cells of the endothelial membrane without crossing it.[122,123] These finding were also transferrable from mouse to CAM model where a transduction of the brain was only observed for AAV9 while the number of genomic copies for AAV2 is in good agreement with the PBS control.

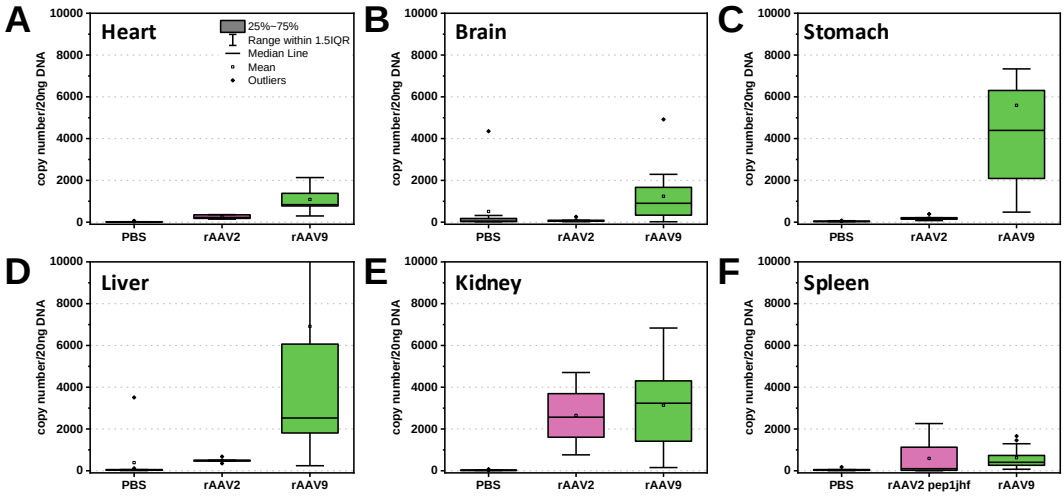


Figure 25: Determination of viral genomes for rAAV2 and rAAV9 via qPCR. A total amount of 1×10^{11} viral genomes was injected systemically and after 48h incubation organs were extracted from chicken embryos and DNA was isolated before qPCR analysis was performed. Box-whisker diagrams were created using Origin2019 and represent the overall replicates for all samples (PBS: 10 replicates; rAAV2:5 replicates, rAAV9: 13 replicates)..

5.2.10. rAAV2 pep1jhf demonstrates efficacy in CAM assays

Another focus lies on the analysis of rAAV with integrated peptide sequences. Larger viral preparations were produced and CAM assays were performed by our collaboration partners. For rAAV2 pep1osy no adverse effects have been observed after systemic injection. In comparison the injection of rAAV2 pep1jhf led to a high mortality of the embryos (7 out of 15) and surviving animals showed severe cerebral haemorrhages and gastric bleeding. The blood clotting time in chicken embryos was shown to be slow and thus, also slight bleedings can be lethal.[124] As two different AAV preparations and furthermore a reduced vector amount (5×10^{10} vectors per egg) were used for these experiments, we can exclude that contaminants or a too high dose led to this severe phenotype.

Rational design of peptides was performed using the PDB model 1IVO which represents the crystal structure of the human EGFR. The chicken EGFR shares about 70% sequence homology with the human EGFR. Taking a closer look into the interaction between the designed peptides and the crystal structure of the human EGFR demonstrates the binding interface with atomic contacts based

on the van der Waals (VDW) radii (Figure 26A, B). Comparing the regions of contacts with the sequence alignment of human and chicken EGFR indicates that interacting regions are almost identical between the organisms (Figure 26C). Thus, transduction not only of the human EGFR in the tumor tissue of A431 cells should be possible, also the transduction of chicken EGFR in the whole organism might be feasible.

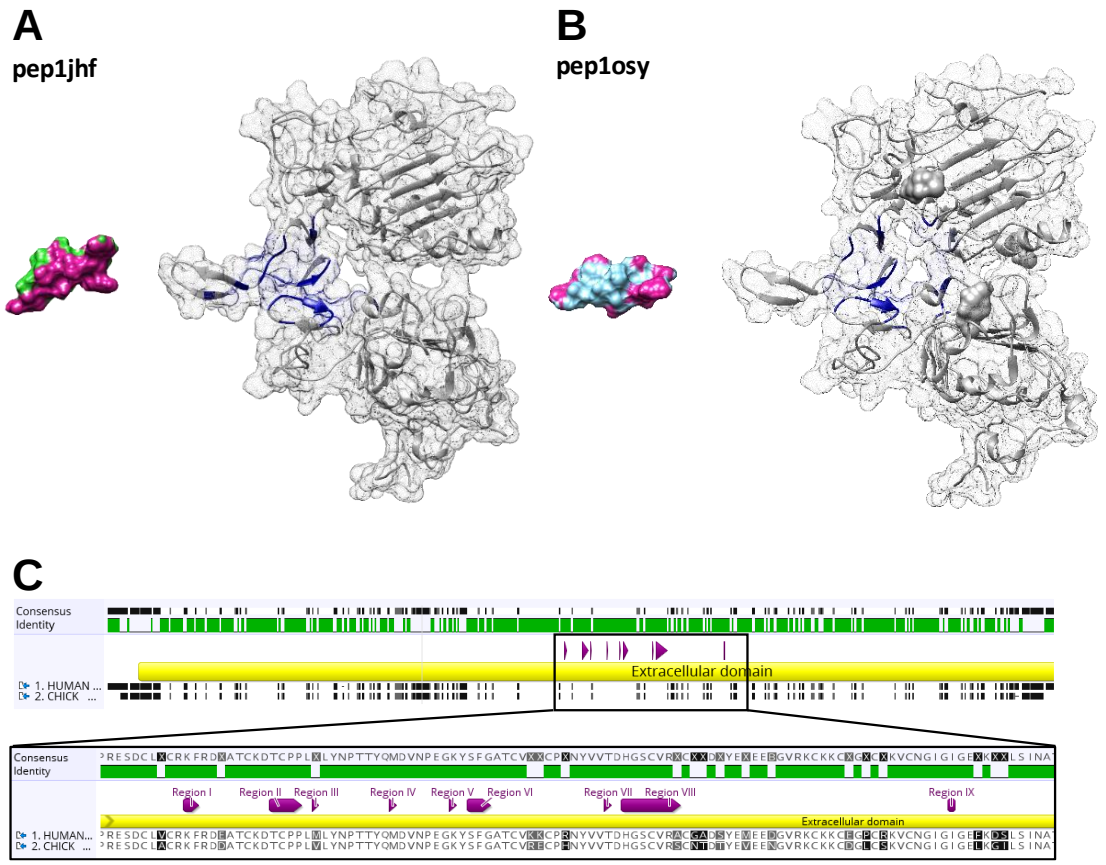


Figure 26: Identification of the binding interface of pep1jhf and pep1osy. PDB files of the bound peptide to the EGFR are the results of the rational peptide design of pep1jhf (A) and pep1osy (B). Contacts between atom pairs of peptide and receptor were determined using UCSF Chimera. A cut-off of 1.0 Å distance was defined to identify atoms pairs. The binding interface is highlighted in magenta (peptide) and dark blue (EGFR). (C) Alignment of the human EGFR (UniProt P00533) and the chicken EGFR (UniProt A0A1D5NZB4) was generated using Geneious. Regions identified from previous contact determination are highlighted in magenta.

Figure 27A-F represents the results of qPCR analysis for detection of delivered genomic copies to the respective organs. Again systemic PBS injection served as negative control and here transduction was not detected. For rAAV2 pep1osy almost no transduction was observed in any tissue. This was not expected as previous cell culture experiments demonstrated that receptor binding, intracellular trafficking, nuclear entry and genome conversion is possible upon integration of the targeting peptide. The most important difference between the previous cell culture experiments and the *in vivo* model lies within the blood clearance and the vascular escape. Since, peptides should be able to interact not only with the human EGFR, but also with the chicken EGFR, rAAV2 variants should also be detected in EGFR-expressing tissues The inefficient transduction of tissues by rAAV2 pep1osy can thus only be explained via the blood clearance and the vascular escape.

The qPCR analysis for rAAV2 pep1jhf showed genomic copies in all organs of the chicken embryo even elevated compared to transduction with the wild-type rAAV2 capsid. The EGFR is widely overexpressed in foetal tissues and thus transduction of all organs is likely. The high lethality after systemic injection due to severe cerebral haemorrhages and gastric bleeding might be explained by qPCR results. The highest amounts of genomic copies were found in these organs. The high amount in the brain is surprising as for the rAAV2 wild-type variant no transduction was observed and it is known that rAAV2 is not able to cross the blood-brain barrier.[123] However, it was shown that the blood-brain barrier of one-day old chicken is humiliated, which might allow for transcytosis of the re-targeted rAAV vector.[125] EGFR expression in the foetal brain is elevated which might contribute to the high transduction efficiency. Gastric bleeding on the other hand was also rarely observed in patients for other EGFR-targeting pharmaceuticals already in clinic.[126,127] Bleeding in the embryo might be the result from an effective EGF receptor blocking. Interruption of the downstream signaling is involved in different cellular processes and if the receptor is not only overexpressed in the tumor but also in the embryo's organs this can have severe side-effects.

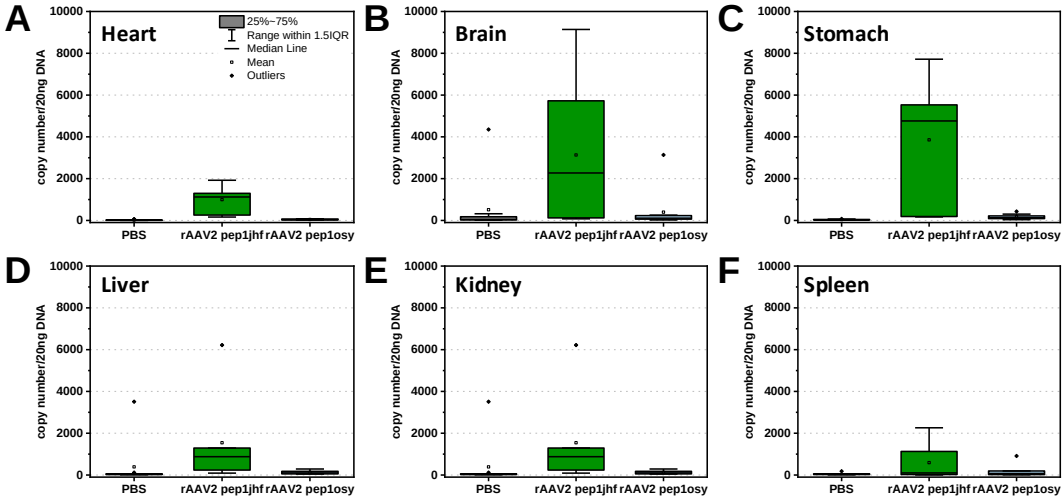


Figure 27: Determination of viral genomes for rAAV2 pep1jhf and rAAV2 pep1osy via qPCR. A total amount of 1×10^{11} viral genomes was injected systemically and after 48h incubation organs were extracted from chicken embryos and DNA was isolated before qPCR analysis was performed. Box-whisker diagrams were created using Origin2019 and represent overall replicates (rAAV2 pep1jhf: 7; rAAV2 pep1osy: 11 replicates)..

Besides analysis of genomic copies in the embryonic organs, the transduction of the tumor tissue is of great interest. The mammalian epidermoid carcinoma cell line A431 was transplanted in the CAM and during the incubation time of 96 h a solid tumor was formed. This model was used to evaluate the vectors properties in an *in vivo* system. Again the genomic copy number in the tumor tissues has been determined in qPCR. A comparison between wild-type rAAV2 and rAAV9 and the re-targeted viral vectors rAAV2 pep1jhf and rAAV2 pep1osy is given in Figure 28. The buffer control as well as wild-type rAAV2 and rAAV2 pep1osy show nearly no transduction at all. rAAV2 wt can only rely on the natural tropism via the primary receptor HSPG. Cell culture experiments showed that in general rAAV2 wild-type is able to transduce A431 cells. In context of the

whole chicken embryo, this vector might show a fast blood clearance and a reduced vascular escape.[120] Results for rAAV2 pep1osy were expected since this vector was not able to transduce any tissue in the chicken embryo. In contrast, rAAV9 wild-type shows slightly elevated copy numbers which can also be related to blood clearance and vascular escape. Particularly remarkable are the values for rAAV2 pep1jhf, which exceed all other values, showing that targeting of the tumor tissues is possible. Comparison of tumor-liver ratios (Figure 28B) showed that the best ratio is obtained for this re-targeted vector.

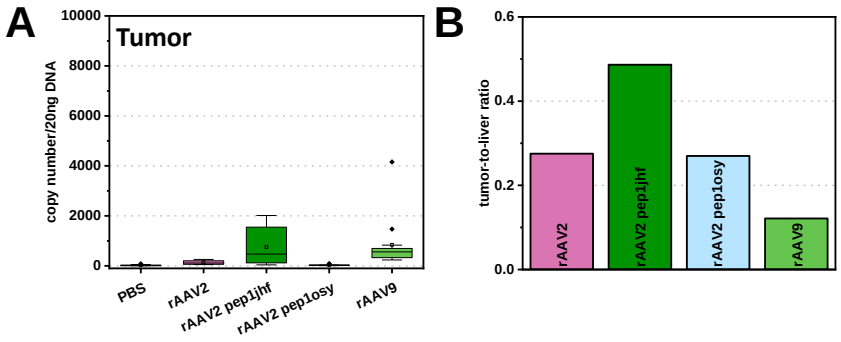


Figure 28: Analysis of tumor tissues transduction via qPCR. (A) The genomic copy number was determined in 20 ng isolated DNA and is presented in a box-whisker diagram. (B) The tumor-to-liver ratio was calculated for all viral vector. The best ratio was obtained for rAAV2 pep1jhf showing that targeting of the tumor tissues is possible via the rational designed peptides.

Summarizing the results from first *in vivo* experiments confirmed the transduction pattern which was already described for rAAV wild-type vectors in mouse experiments. These findings provide necessary information to avoid more time-consuming experiments in mice. Experiments conducted for the re-targeted rAAV variants showed that the transfer from an *in vitro* cell culture-based system to an *in vivo* model can be complex. Here, the transfer from cell culture to *in ovo* experiments was not successful for rAAV2 pep1osy but showed interesting results for rAAV2 pep1jhf. rAAV2 pep1jhf in general shows an elevated transduction efficiency. This was also observed for transduction of the tumor tissue. The severe complications after virus injection show the need for further optimization of the re-targeting viral vector. The side effects need to be reduced in the future to allows for a specific transduction of the tumor tissue. Combining the pep1jhf capsid with an intracellular transcriptional or translational targeting might allow for a more specific expression of the gene of interest. Being under the control of, e.g. a tumor specific promoter, might prevent gene expression in healthy tissue and promotes expression in the tumor tissue.

5.3. Establishing plasmid systems for the production of different AAV serotypes

Results of this project were partially summarized in manuscripts. Results regarding AAV2 have been submitted to Scientific Reports with the title “rAAV engineering for capsid-protein enzyme insertions and mosaicism reveals resilience to mutational, structural and thermal perturbations”.

5.3.1. Characterization of AAV2 produced with the virus construction kit

Production of rAAVs using the helper-free plasmid system requires three plasmids. The helper plasmid delivering necessary adenoviral elements is accompanied by ITR and RepCap plasmid. The RepCap (pZMB0216_Rep_VP123_453_587wt_ p5tataless) plasmid used as a starting point in my work was developed during the iGEM competition 2010 in Freiburg.[128] It offers the possibility to easily introduce modifications at residue position 453 and 587 located within the variable regions IV and VIII of the VP protein. Genetic information that is finally packaged in rAAV capsids is provided by the ITR plasmid. We constructed an ITR plasmid, in which the viral ITRs are part of a pUC19-based backbone (pZMB0522_ITR_EXS_CMV_mVenus_hGHpA). Construction of a new ITR plasmid was necessary as with the iGEM ITR plasmid AAV production was not possible any more. Recombination events during the propagation of *E. coli* led to larger deletions within the important ITR structures, that are necessary during the production process. Verifying the correctness of the ITR sequences is challenging due to the strong secondary structure formation that defies Sanger cycle sequencing. A protocol was established by Julian Teschner that prior to the sequencing reaction includes a restriction digest with BsaHI. This enzyme generated two halves of the ITRs that are each on their own suitable for sequencing. We observed a 11 bp deletion in the 3’-ITR while the 5’-ITR is completely intact. First we assayed general functionality of the ITR plasmid in combination with either the RepCap plasmid (pZMB0216_Rep_VP123_453_587wt_ p5tataless) or a commercially available counterpart (pAAV-RC, GenBank: AF369963.1) in small-scale transfections. Genomic titers were determined for the crude cell lysates in qPCR reactions of several biological and technical replicates. As seen in Figure 29A no difference can be observed.

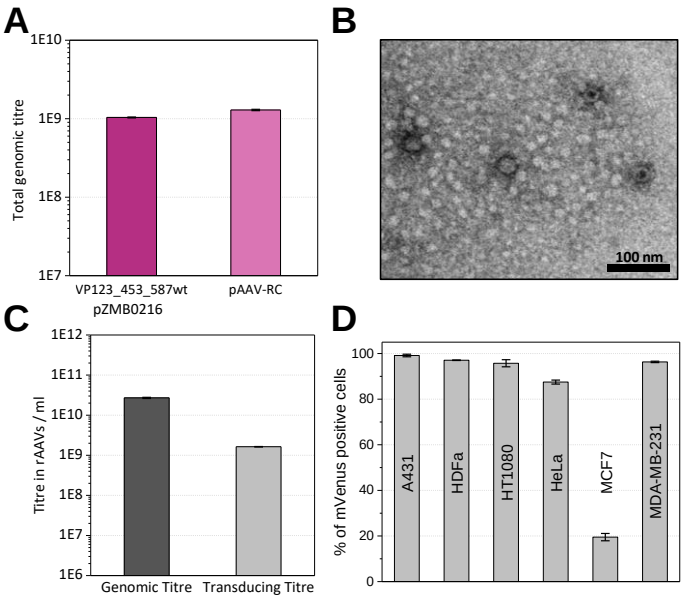


Figure 29: Characterization of rAAV2 after production using the helper-free plasmid system. (A) Comparison of genomic titers from crude cell lysates using RepCap (pZMB0216) and commercial pAAV-RC. Standard deviations of three biological and two technical replicates were calculated for each sample type in this comparison. (B) Transmission electron microscopy of a rAAV2 preparation at a magnification of 39,000. Viral particles were detected with a size of about 25 nm. (C) Analysis of a larger purified rAAV2 preparation for genomic and transducing titer using qPCR or transduction of HT1080 cells respectively. (D) Transduction of different cell lines with iodixanol-purified rAAV2 at a MOI of 10,000. Successful transduction was visible upon expression of the gene of interest mVenus which was measured by counting 10,000 events for two biological duplicates using a FACScalibur. Data analysis was performed using FlowJo V10.

In a next step, a larger production of rAAV2 with this plasmid system was generated and purified via a discontinuous iodixanol gradient ultracentrifugation. The protocol for this purification method was established in this work and is based on methods described previously.[129] Before this ultracentrifugation protocol was established purification was carried out using a protocol published by Guo et al.[130] Viral preparations purified using this precipitation-based method were used for transmission electron microscopy. In transmission electron micrographs, capsid diameter measurements resulted in an average of 23.8 ± 1.8 nm, which is in good agreement with the expected value of 25 nm (Figure 29B).[131] Manually counting over 500 particles yielded a proportion of full capsid between 60% and 80%. A main problem of the precipitation-based purification method was observed especially during transduction experiments. The unsatisfying purity of the preparation led to a strong reduction in cell viability and transduction efficiency, which required a change in the method. The first ultracentrifugation protocol was based on caesium chloride as density gradient medium in a continuous gradient. The time-consuming protocol was very quickly replaced by the density gradient medium iodixanol.[129] Genomic titers of viral preparations after iodixanol ultracentrifugation were determined using qPCR and are given in Figure 29C. The ratio of genomic copies to infectious units yields the specific infectivity of the preparation, which was 16:1. This is in agreement to reported results as for wild-type AAV a ratio of 1:1 and for rAAV2 a ratio between 55:1 and 124:1 have been observed.[132] Functionality of the rAAVs was investigated in transduction assays on HT1080 cells. These cells are known to express high levels of the rAAV2s primary receptor

HSPG.[38,39] Successful transduction was detected by the expression of the delivered fluorescence reporter mVenus using flow cytometry. In addition to commonly used target cell lines such as HT1080 and Hela, we treated the cancer cell lines A431, MCF7, MDA-MB-231 and the normal cell line HDFa with a MOI of 10,000. Figure 29D presents the results of a flow cytometry analysis, showing that our rAAV2 was able to transduce a variety of different cells with high efficiencies. In agreement to previous reports, only the breast cancer cell line MCF7 showed lower transduction values.[112]

Production of mosaic viral particles was described for our plasmid system for *N*-terminal VP2 protein fusions.[60] A mutation in the RepCap plasmid leads to knock out of the VP2 start codon (pZMB0299). An *N*-terminally-modified VP2 sequence is provided *in trans* on a separate plasmid. Splicing during VP protein expression results in two mRNA transcripts that code for VP1 and for VP2/3, respectively, and a leaky-scanning mechanism induces expression of either VP2 or VP3.[13,14] Thus, when the VP2 sequence is extracted, the coding sequence of VP3 is always included. For *N*-terminal modifications this does not pose any problems since VP2 and VP3 share the same *C*-terminal sequence but differ in their translation start and thus their *N*-terminus. Since the 587 loop region lies within the coding sequence of both VP proteins, VP2 modifications require that the parallel expression of a likewise modified VP3 has to be suppressed.

We saw the need for a detailed expression analysis of the plasmid system. The starting point for our experiment was the CMV VP23 plasmid from literature (CMV VP23).[60] As expected, a strong expression of VP2 and VP3 proteins is observed (Figure 30A, lane 2). In comparison, expression of all three VP proteins after transfection with the unmodified RepCap plasmid shows the expected molar ratio between the three VP proteins (1:1:10) (Figure 30A, lane 1). To prevent undesired VP3 expression, the VP3 start codon knocked out by a point mutation in the start codon (ATG to ATC (Ile)) in CMV VP2 (Figure 30A). As seen in the third lane of Figure 30A, expression of VP3 is still observed, which might be due to a second start codon located 24 bp downstream. To suppress leaky scanning, a strong Kozak sequence (GCC ACC) was introduced into CMV VP2 in front of the start codon. Finally, solely the expression of VP2 (lane four of Figure 30A) was detected with an expected increase in chemiluminescence intensity, indicating a higher level of expression.

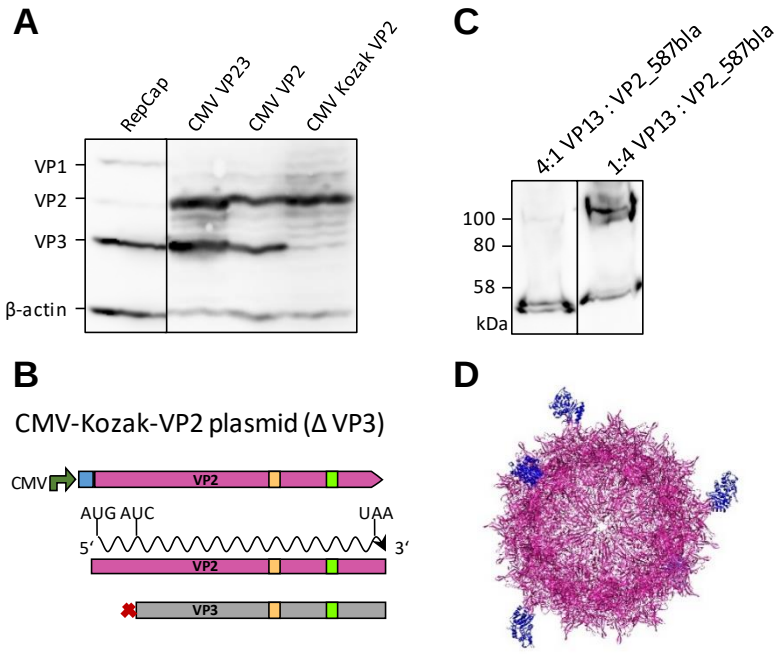


Figure 30: Development of a mosaic rAAV system allowing for partial modification of the 453 and 587 loop region. (A) Expression analysis of VP proteins after transient transfection of CMV VP plasmids. For the RepCap plasmid VP protein expression was confirmed by the anti-VP antibody B1 with the expected ratio of 1:1:10 for VP1, VP2 and VP3. Three versions of a VP2 plasmid were analyzed: 1) VP2 and VP3 expression cassette under CMV promoter control (lane 2, CMV VP23), 2) VP2 expression cassette with VP3 start codon knock out (lane 3, CMV VP2), and 3) VP2 expression cassette with upstream Kozak sequence and VP3 start knock out (lane 4, CMV Kozak VP2). (B) Schematic explanation of the final CMV-Kozak-VP2 plasmid. (C) Western Blot analysis of crude HEK293 lysate after quadruple transfection using different plasmid ratios. The ratio of the Cap delivering plasmids was tested in two approaches either 1:4 or 4:1 (VP13:VP2_587bla) showing that a reduced amount of the CMV-containing plasmids also reduces the expression of the VP2-587bla. (D) Crystal structure representation of a mosaic rAAV2_VP2_587_bla with incorporation of five β-lactamase enzymes. The structure was assembled from PDB 1LP3 and PDB 3DTM using UCSF Chimera.

The gene of the stabilized β-lactamase variant 14FM was cloned into CMV Kozak VP2. The resulting plasmid was used for rAAV production in combination with pZMB0299, the ITR plasmid and pHelper. Different from the three-plasmid system, the molar ratio of these plasmids needs to be optimized for transfection. The two plasmids contributing to the expression of VP proteins need to be in an optimal proportion to provide the right amount of each VP protein (1:1:10) for correct capsid assembly. In this experiment it was observed that the molar ratio 5:5:1:4 of pHelper:ITR:Rep2Cap_VP13:CMV_VP2_587bla was associated with a greatly increased proportion of VP2 protein (Figure 30C). This would lead to a higher portion of modified VP2 proteins in the assembled viral capsid compared to the ratio of 1:1:10 of VP1, VP2 and VP3 in the wild-type AAV. Due to the differences in promoter strength (AAV's p40 is much weaker compared to CMV) we used a change in plasmid ratio to reduce the expression level of VP2_587_bla and thus the resulting capsid modification level. In addition to Western blot analysis, viral samples were also analyzed regarding their genomic titer. Determination in samples of crude cell lysate showed that reduced expression VP2_587_bla protein results in roughly three times higher amounts of viral particles (7.98×10^{10} vg/ml for a 4:1 VP13:VP2_587_bla ratio compared to 2.48×10^{10} vg/ml for a 1:4 VP13:VP2_587_bla ratio).

Shown to yield more promising results regarding the number of total modified VP2 proteins, mosaic rAAV2-VP2_587bla particles were produced using the 5:5:4:1 ratio of pHelper:ITR:Rep2Cap_VP13:CMV_VP2_587bla. We assume that the purified particles present a β-lactamase in every VP2 protein, thus a total of five modified proteins as illustrated in Figure 30D. Incubation of HT1080 cells with a MOI of 50,000 resulted in about 57 ± 2% mVenus positive cells. Comparison with transduction values of rAAV2 wt shows a significant reduction of transduction ability for the enzyme-bearing particles.

5.3.2. Transferring the virus construction kit to rAAV serotype 6 and 9

A broad range of AAV serotypes and variants is known and for some even a more detailed characterization was performed. In this work the previously designed virus construction kit for AAV2 was transferred to other serotypes to promote easy modification and further characterization of AAV6 and AAV9. For both serotypes integration into variable region VIII was already proven for the integration of peptide sequences.[61,62] The integration sites in both Cap reading frames was identified at position 587 and furthermore also the tolerability of insertion in variable region IV at residue position 453 was to be assayed (Figure 31A). The system developed for rAAV2 offers the possibility for integration of peptide motifs via single cutting restriction sites that were introduced under the premise that the coding region is unaltered. Here a combination of SalI and SspI as well as BamHI and PvuII allow for integration of DNA sequences with corresponding overhangs in the 453 region and the 587 region respectively. Applying the same restriction site combinations directly to the sequence of AAV6 and AAV9 was not possible and therefore we opted to find other single cutting enzymes in this region. A detailed sequences analysis showed that for both serotypes single cutting enzymes could be found in the 453 and the 587 region (Figure 31B). A combination of BspEI and BsrGI was chosen to make the 453 region accessible, while MscI and BamHI were chosen for the 587 region. The pSB1C3_001 backbone needed to be adapted to the new cloning strategy as BspEI and MscI were detected in the chloramphenicol acetyltransferase gene of the backbone (Figure 31C). The new backbone pSB1C3_002 was generated by site-directed mutagenesis and was further used as a backbone for AAV6 and AAV9 *cap* sequences. The coding sequences of Cap6 and Cap9 have each been ordered as a string synthesis and were recloned in the new pSB1C3_002_Rep2_Cap2 plasmid via the unique restriction sites SwaI and PstI. Final constructs were analyzed for their correctness using Sanger cycle sequencing.

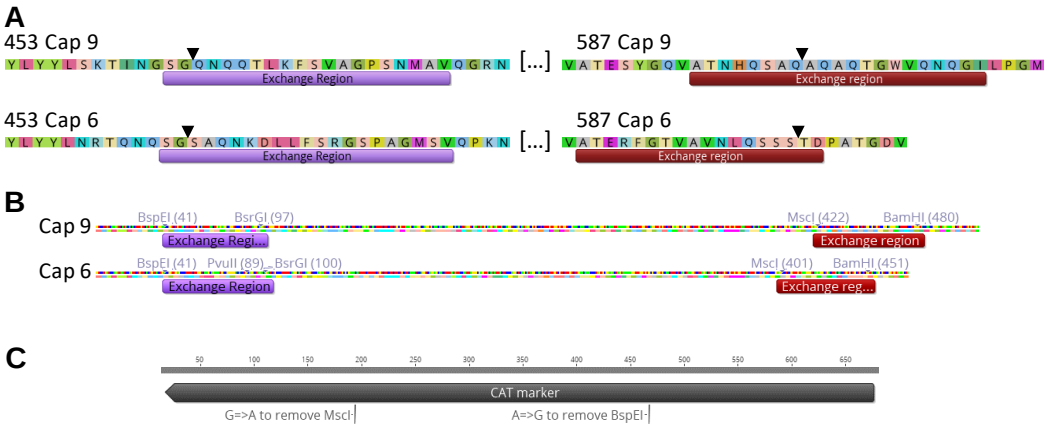


Figure 31: Establishment of a cloning strategy to allow for easy accessibility of the variable regions IV and VIII. (A) Amino acid sequence around the 453-(violet) and 587-exchange region (dark red) highlighting the insertion site at position 453 and position 587 respectively for AAV9 and AAV6. (B) DNA sequence of both exchange regions for AAV9 and AAV6 showing the potential single cutting restrictions sites. (C) Chloramphenicol acetyltransferase gene (CAT marker) of the pSB1C3_002 backbone eliminating the MscI and BspEI recognition sites by site-directed mutagenesis.

5.3.3. Characterization of AAV6 and AAV9

Functionality of AAV6 and AAV9 plasmids in rAAV production was tested in adherent HEK293 cells. Viral preparations of AAV6 and AAV9 with a wild-type capsid were used for a first characterization of the serotype. A Western blot analysis was performed with the anti-VP antibody B1 to prove the correct size and ratio of all three VP proteins. As seen in Figure 32A visible bands are detected with the expected size and the frequency of the VP3 protein is as expected strongly elevated compared to VP1 and VP2 emphasizing the 1:1:10 ratio. For AAV9 a more detailed analysis was performed using transmission electron microscopy and atomic force microscopy. Figure 32B presents AAV9 particles after staining with uranyl acetate. The size of the presented particles was determined to be 25.23±0.94 nm. In AFM micrographs (Figure 32C,D) the diameter was calculated with 33.43±1.00 nm. The method of determination differs between both methods and is dependent on the analysis program, which might explain such great difference between the diameter measurements of viral particles of the same preparation.

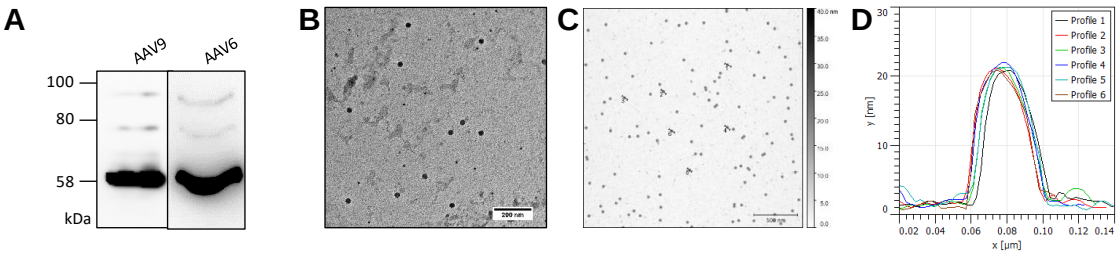


Figure 32: Characterization of new rAAV variants. (A) Western blot analysis of rAAV6 and rAAV9. VP proteins were detected upon incubation with the B1 antibody (Progen). (B) Transmission electron micrograph of rAAV9 after negative staining with uranyl acetate (2%). Images were acquired with a Philips CM100 (PW6021) at a magnification of 21.000×. Analysis was performed with ImageJ (C, D) Atomic force microscopy (AFM) was performed on a Multimode 8 AFM (Bruker) with Tap300Al-G cantilevers in tapping mode in air. Data analysis was performed using Gwyddion 2.48.

A direct comparison of the self-cloned Rep2Cap9 plasmid (pZMB0551) and an AAV9 version distributed via the Penn Vector Core (PVC) (pZMB0504) was carried out and surprisingly a great difference in genomic titer was observed. Production with the PVC Rep2Cap9 plasmid resulted in average in genomic titers one potency higher compared to our Rep2CapX plasmids. First experiments were conducted to see which part of the plasmid is accountable for such a great difference. Both plasmids have almost identical sequences for Rep and Cap protein but differ in their backbone. As previous experiments comparing Rep2Cap2 (pZMB0216) with the Agilent variant (Figure 29A) resulted in equal titer we thought the backbone might not be relevant and took a closer look on the Rep and Cap part. Here we saw differences in the *N*-terminal sequences of the *rep* coding sequences. As several alternative splicing occurs at this site, we hypothesized that changing the *N*-terminal sequence would lead to an increase in titer. The resulting variant pZMB0576 was tested in adherent HEK293 and suspension 293F cells. Results of this experiment are presented in Figure 33 and prove that, regardless of the production cell line, the difference of one potency is still observed. Thus, reasons for this difference in production efficiency still need to be analyzed and might unexpectedly be related to the different plasmid backbone.

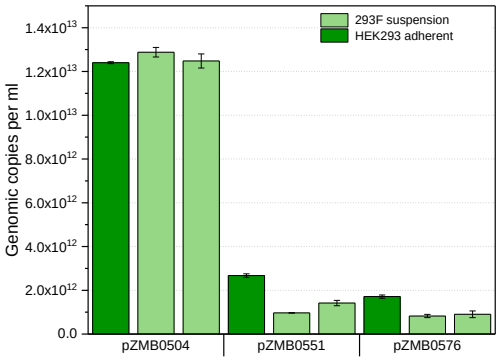


Figure 33:Comparison of different AAV9 RepCap plasmids. Triple transfection were performed with the same pHelper and ITR plasmid (pZMB0522) but using either pZMB0504 (PennVectorCore), pZMB0551 (Rep2Cap9 from string synthesis) or pZMB0576 (Rep2Cap9 after exchange of the N-terminal Rep part). Genomic titers were determined from crude cell lysate using a standard curve equilibrated with the same amount of lysed untransfected cells.

Furthermore, both serotypes were analyzed for their temperature stability using a method based on genome release upon thermal incubation. Subsequent qPCR analysis results in a sigmoidal relationship between incubation temperature and percentage of genomic copies which can be used to determine the disintegration temperature T_d. Various methods have been described, such as differential scanning calorimetry (DSC), differential scanning fluorimetry (DSF) and electron microscopy, all of which monitor capsid disintegration but do not detect DNA release.[133–135] We assume that rAAV particle integrity can be analyzed by DNase accessibility and repurposed the standard assay to determine DNaseI-resistant viral particles. Results for the thermal stability of AAV9 and AAV6 are presented in Figure 34A and B respectively. Regarding this disintegration temperature AAV9 seems to be more stable compared to AAV6. In comparison of AAV6 to AAV2, the disintegration temperatures are located in the same range (T_d(AAV2)=56.1 °C), T_d(AAV6)=56.9 °C

(Manuscript under revision)), while AAV9 seems to be more stable regarding the release of genomic DNA (T_d(AAV9)=61.4 °C). Comparing these values with data obtained by DSF; DSC or electron microscopy shows that overall values are 15-20 °C lower. Capsids appear to be more stable in terms of protein stability than protection of their genomic content. The identified stability trends are not fully reflected in the protein stability data. Here, AAV6 and AAV9 should have approximately the same stability, which is almost 10 °C higher than serotype 2.[134] However, it can be seen that AAV9 has a considerably higher disintegration temperature than AAV6 and AAV2. A model of the dependency between genomic and protein stability for AAV2, AAV6 and AAV9 is given in Figure 34C.

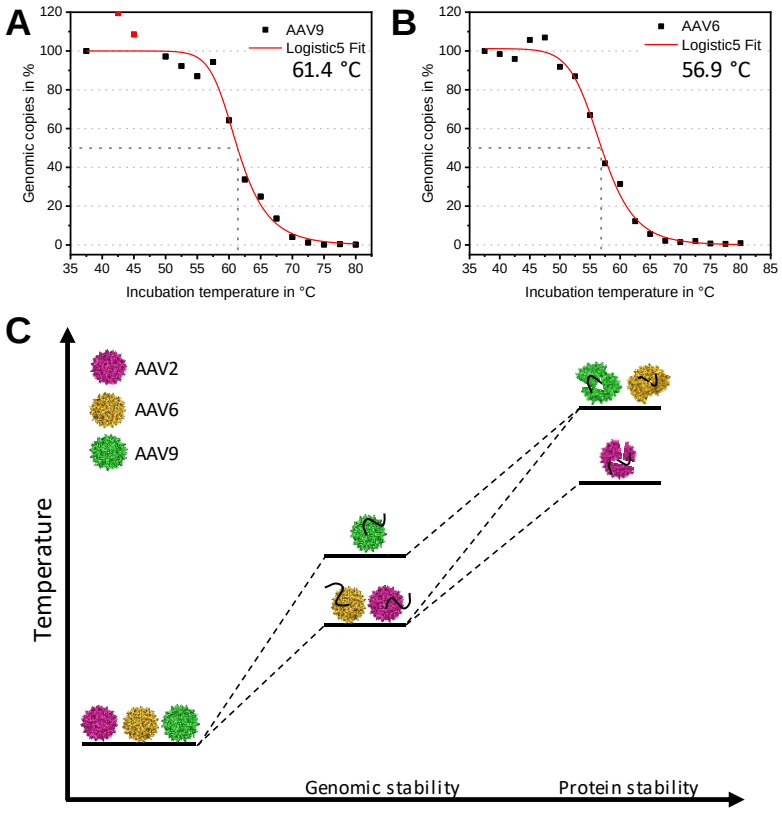


Figure 34: Characterization of AAV6 and AAV9 stability and transduction ability. (A, B) Thermal stability assays of AAV6 and AAV9 measured in PBS by qPCR. The percentage of intact genomic copies is plotted against the incubation temperature in °C. Each point represents the standard deviation of a technical duplicate. Fitting curves (red) were calculated using a logistic5 function to determine the disintegration temperature in Origin2019. Red highlighted data points were masked and have not been integrated into the fit. Temperature are given within the graph. (C) Degradation of AAV capsids upon thermal heating. This model described the relationship between genomic and protein stability that is evaluated using different methods. Data from literature is compared to the data from A and B.

Three different AAV serotypes are now available for further analysis. Transduction ability of these wild-type variants in general was described in literature for a broad range of cell lines.[112] Not all cell lines we use in our laboratory have been included in this overview and thus, we were interested how the different serotype behave in context of these cell lines. A broad range of cancer cell lines including A431, A431 AAVRko, MCF7, MDA-MB-231, HeLa and HT1080 beside the normal cell human dermal fibroblast cell lines (HDFa) have been tested using a MOI of 10,000. Cells were

incubated and analyzed for the delivered fluorescent protein mVenus by flow cytometry. Results are presented in a heat map diagram in Figure 34C and are in good correspondence with in literature described values.[112] For AAV2 and AAV6 with a wild-type capsid high transduction efficiencies were observed in nearly all types of cells. Low transduction efficiencies for A431 AAVRko were anticipated as all serotypes (AAV2, AAV6 and AAV9) are dependent on this receptor.[40] The overall low transduction ability of AAV9 was expected as for *in vitro* cell culture experiments this serotype was already described to be ineffective. Interestingly, these observations cannot be made in *in vivo* mouse experiments, where AAV9 was able to reach maximal vector expression.[119]

MOI 10,000	AAV2	AAV9	AAV6
A431	99	3	94
A431 AAVRko	14	2	10
MCF7	20	3	51
MDA-MB-231	96	8	85
HDFa	97	2	59
HeLa	88	2	92
HT1080	99	43	99

Figure 35: Characterization of transduction ability. AAV2, AAV9 and AAV6 have been used to transduce a moiety of different cell lines with a MOI of 10,000. Cells were analyzed for their mVenus expression after 48 h incubation. 10,000 events were counted using a FACScalibur system and data was analyzed using FlowJo and Excel.

To study the effects of motif insertion on the natural serotype tropism a hexahistidin-tag has been inserted into 588 position of AAV6 and AAV9. Western blot analysis of AAV9 variants showed the presence of the VP3 protein in the viral preparation. Due to low sample concentration VP2 and VP1 are not detectable.

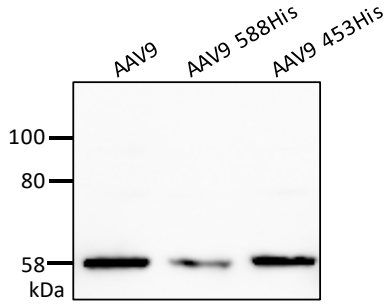


Figure 36: Western blot analysis of AAV9 variants with hexahistidin-tags in 588 or 453 position, respectively in comparison to AAV9 wild-type. VP proteins were detected upon incubation with the B1 antibody (Progen) and chem-iluminescence detection of the secondary antibody.

For AAV9 transduction experiments were conducted on HT1080 cells, as this was the only cell line where transduction was observed in higher efficiencies. Here, the insertion of the hexahistidin-tag led to a total decrease of transduction ability down to below 1%. Data for AAV9 were already presented in Figure 23 and also demonstrate a strong reduction of transduction ability from 99% to

about 16% on both MCF7 and A431 cells. Both experiments show the importance of the loop structure for the tropism. K531 was described as the determinant for AAV6 HSPG binding and is located in the three-fold spike.[136] Binding of AAV9 to N-terminal galactose was characterized by N272 and W503, which are also located in the three-fold spike of the capsid.[37] In both cases the integrated sequence might be partially shield the important amino acids from the elongated loop structure.

5.4. Optimizing the production of rAAVs

Projects described in this chapter are summarizing different attempts to optimize the production of rAAVs in mammalian cells.

5.4.1. Production of rAAVs using a 293F suspension cell line

The results of this project were presented at the 25th European Society for Animal Cell Technology Meeting in Lausanne 2017 and published in BMC Proceedings with the title "AAV production in suspension: evaluation of different cell culture media and scale-up potential". The original publication is included in the appendix.

Production of rAAVs using mammalian cells mostly relies on the use of adherent HEK293 cells.[137] Here, transfection protocols as well as downstream processes yield in preparations with high titer and high purity. For upscaling processes the use of adherent cells is problematic since the growth area of cells is limited. Suspension cultures on the other hand allow for higher cell densities and require less space for cultivation. However, production and purification methods need to be adapted to this method. We started with experiments on the commercially available cell line 293F (ThermoFisher Scientific), which was derived from the adherent HEK293 cell line. The calcium phosphate transfection protocol, which is routinely used for adherent HEK293 cannot be transferred to suspension culture. Thus the first step to bring rAAV production to suspension was establishment of a suitable transfection protocol. Since expertise in working with suspension culture was available in our group for CHO K1 cells, we adapted the polyethylenimine (PEI) transfection protocol from this cell line. The transfection efficiency was determined from the ITR plasmid harboring the fluorescent reporter mVenus under control of a CMV promoter. Different DNA:PEI ratios in combination with different DNA amounts per 3×10^6 cells were tested to increase the overall efficiency and best results were reproducibly obtained for 2 µg DNA in a 1:4 DNA:PEI ratio. Transferring the transfection protocol from 6-well plates to shaker flasks even increased the transfection efficiency (Figure 37).

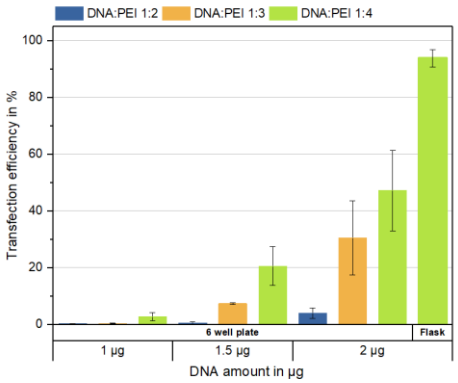


Figure 37: Transfection efficiencies were determined by mVenus expression 48 h after triple-transfection using flow cytometry analysis. The DNA amount is given per ml of culture volume (3×10^6 cells/ml starting culture).

These first pre-experiments were used to analyze optimal conditions for rAAV production in 293F suspension culture. Using triples transfections in shaker flasks allowed for investigating different parameters each day to determine the optimal time point for harvest of the culture (Figure 38A). As expected the viable cell density (VCD) is increasing over the production process with stable high viabilities. The amount of DNaseI-resistant particles was determined from raw cell lysates using qPCR. It was shown that the highest titer was typically achieved between day three and four. A decrease in viability marked the decline in genomic copies per ml, showing that a prolongation of the process, e.g. by addition of a feed, would probably not increase yield. This experiment also clearly proved that in downstream processes the focus needs to be laid on purification from the cell pellet, as nearly no viral particles were observed in the cell culture supernatant.

In a first scale-up process, the rAAV production was transferred to a 2 l bioreactor (Figure 38B). This scale-up showed potential bottlenecks, because the transfection protocol that was applied required several steps that are not easily realized in bioreactor processes. Prior to the bioreactor process it was observed that a medium change before transfection significantly increases the transfection efficiency. To prevent the centrifugation step which would increase the risk of contamination, the culture was grown to high cell densities and finally diluted to 3×10^6 cells/ml prior to transfection. Transfection efficiencies in bioreactors of up to 55% were comparable to that obtained in a simultaneous shaker flask experiment. Transfection efficiencies were lower compared to prior experiments due to the change in the transfection protocol and controlled conditions in the bioreactor. Nonetheless the titer with up to 1×10^5 genomic copies per cell was elevated compared to that of shaker flasks.

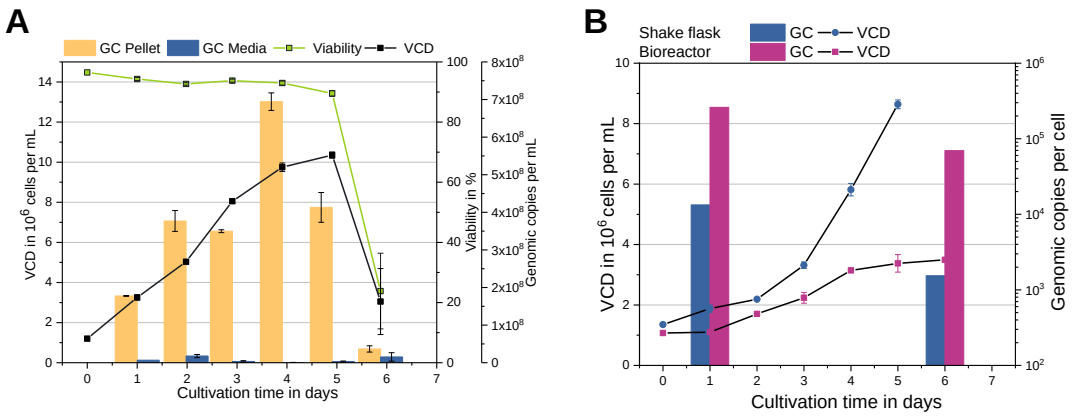


Figure 38: Transferring rAAV production to suspension 293F culture. (A) Viable cell densities (VCD), viabilities and genomic copies per ml (GC) of a rAAV production with 293-F batch cultivations in shaker flasks. Genomic copies per ml refer to the titer determined in 1 ml culture volume. Error bars represent biological and technical duplicate measurements of samples. (B) Viable cell densities and genomic copies per cell of a rAAV production with 293-F batch cultivation in a 2 l bioreactor. For reasons of comparability between shaker and bioreactor data genomic copies are given per cell. Error bars represent technical duplicate measurements of samples.

First experiments with 293F cells in HEK-TF medium showed promising results of transferring rAAV production from the adherent system to suspension. After improvement of transfections by the adjustment of DNA amounts in small scale experiments, AAV production was analyzed in shaker flasks. The batch process showed an expected increase in cell density with low variability between biological replicates. The genomic titer increased according to the viable cell density until day four where a sudden drop started. From optimized protocols, a batch process in a 2 l bioreactor was carried out. Interestingly the bioreactor cultivation resulted in lower overall viable cell densities but in higher genomic copies per cell compared to shaker flasks. These results are comparable to already published data for suspension cells. The work on rAAV production in suspension 293F cells was progressed in the master theses of Irina Schierbaum and Thilo Pohle. Until now we were able to produce comparable yields of rAAVs in suspension culture than in adherent cells. The main problem which was not solved during this work, was the downstream processing yielding in rAAV preparations of high yield and purity. Standard purification protocols used during this thesis were relying on gradient density ultracentrifugation with iodixanol. Transferring this protocol from purification from adherent cells was not successful and we think that this is related to the composition of the media which remains unknown. Probably ingredients of the medium influence the purification and hinder the viral particles to sediment on the gradient. Thus the viral vectors used in this thesis were produced in adherent HEK293 cells.

5.4.2. Characterization of a novel rAAV production cell line

Results of this project were summarized in a manuscript with the title “HEK293-KARE, a cell line with stably integrated adenovirus helper sequences simplifies rAAV production” and submitted to BMC Biotechnology.

HEK293 cells were generated by transfection with fragments of mechanically sheared adenovirus 5 DNA. During this process the adenoviral elements E1A and E1B were randomly integrated into chromosome 19 (19q13.2) and in previous work it was shown that both gene products are expressed.[138–140] Production of rAAV in mammalian HEK293 cells requires additional helper functionality delivered by the pHelper plasmid.[47,141] This plasmid unites the necessary adenoviral elements E2A, E4 and VA RNA and has a size of about 11.6 kb. For triple-transfection large quantities of pHelper plasmids are required beside the RepCap and ITR plasmid. In collaboration with Kathrin Teschner and Julian Teschner a cell line was planned with stably integrated pHelper sequences. In a first step the resistance gene of a blasticidin deaminase (*bsd*) was cloned into the pHelper plasmid to allow for selective pressure during the cell line generation. HEK293 cells were transfected with the linearized plasmid and after three days, selective pressure was applied. Two more weeks of incubation were necessary to detect cell growth and subsequently a limiting dilution was performed to select for single cell colonies. Four monoclonal cell clones HEK-KARE1a-d were isolated and further characterized for successful integration. Resistance to blasticidin already proved that an integration possibly occurred because any remaining plasmid contaminants had been extremely diluted during long time period of incubation. To prove genomic integration, we extracted genomic DNA using Chelex100 and performed PCR reactions with two sets of primers to prove the fully integration of the large fragment. Bsd-for and Bsd-rev as well as Ad5-for and Ad5-rev anneal at different at both ends of the plasmid fragment and in the case of a successful complete integration, amplification products with a size of 499 bp and 524 bp respectively for E2A and *bsd* should emerge. Experimentally obtained fragments are shown in Figure 39A for all samples, demonstrating that both genes were integrated successfully into the genome of HEK293. Microscopy analysis of clone KARE1c proved in comparison with the parental HEK293 cells line that cell morphology was not affected by the integration of such a large fragment (Figure 39B). Furthermore, cell growth was monitored over a time period of 70h and the doubling time of each clone was calculated in comparison with HEK293 cells. Here, data points indicate that growth was not strongly altered upon integration and doubling times for clone HEK293-KARE1a and 1c were equally compared to HEK293 (Figure 39C). A slower growth rate was fitted for HEK293-KARE1b, which can be explained by the higher initial cell density limiting growth after 96 h.

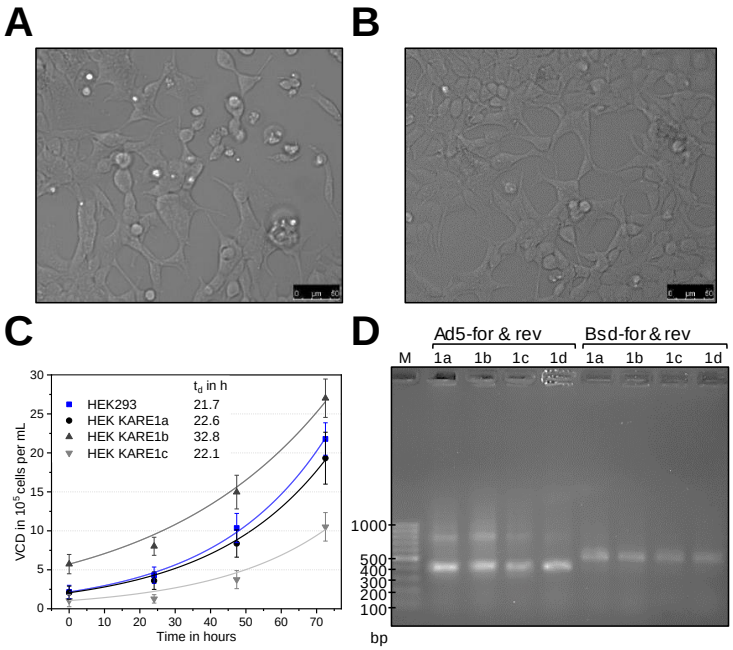


Figure 39: Analysis of cell morphology and growth. (A, B) Comparison of cell morphology in bright field images of (A) HEK293-KARE1c and (B) HEK293. The morphology of HEK293-KARE1c looks similar to the one of the parental HEK293 cell line. (C) Growth curves of HEK293-KARE1 clones and HEK293 were recorded. Cells were seeded on a 12-well plate and harvested after the indicated incubation time. Viable cell density and viabilities were analyzed using an automated cell counting system (CEDEX, Roche Diagnostics). Doubling times t_d were calculated from an exponential fit with Origin 2018 (OriginLab) and are given in the inset of the graph. (D) Agarose gel of the PCR analysis of pHelper-BSD integration. Genomic DNA of the four clones HEK293-KARE1a to HEK293-KARE1d probed by PCR using the primer pairs given above each lane.

For successful rAAV production gene products of all incorporated adenoviral elements needs to be present. The expression level of E2A and E4 gene products was verified after extraction of mRNA from the cells. RT-qPCR analysis was performed after cDNA synthesis with three different subsets of primers for HEK293 cells and clones Kare1a-c. Expression of E1A was used to normalize between the parental cell line and KARE1 clones as expression levels should be equal. The primer pair for E2A expression analysis anneals at the coding sequence of the DNA-binding protein of human adenovirus type 2.[142] E4 expression was verified using a primer pair annealing at open reading frame 6 (ORF6) of E4 generating the 34k protein, which is mainly involved in the AAV production.[143] Figure 40A shows the elevated relative expression level of both gene products after calculation of the ΔΔCt value in the KARE1 clone compared to the parental cell line. Furthermore, the qPCR amplicons were analyzed for their correct size in an agarose gel (Figure 40B). For all qPCR products the corrects size is observed with 68 bp, 65 bp and 143 bp for E1A, E2A and E4 respectively. Interestingly, the expression level of E2A was found to be higher in all cases in comparison to E4. Each early adenoviral gene has its own promoter which becomes active upon expression of E1A.[144] The strength of expression is dependent on this individual promoter, which can lead to totally different amounts of mRNA even if integration occurs at the same genomic site. A low relative expression level was expected for all surviving clones as it was previously described, that high E2A expression would compromise cell growth in the presence of E1 proteins.[145] The

gene dose necessary for high-titer rAAV production is to our knowledge unknown. But we suggest that even low levels of all necessary adenoviral elements is sufficient for rAAV production.

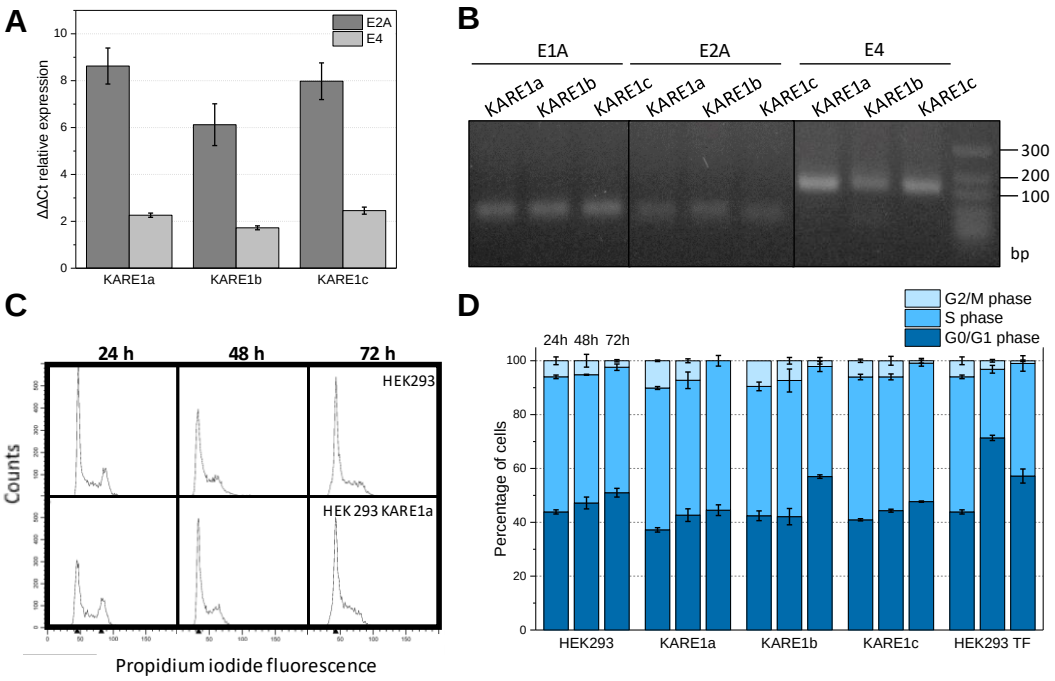


Figure 40: Characterization of HEK293-KARE1 cell clones. (A) RT-qPCR analysis of mRNA extraction of HEK293 and KARE1 cells. The error bars represent the standard deviation from three technical replicates. ΔΔCt values were calculated from the difference between expression in HEK293 and KARE1 cells after normalizing to the E1A gene expression. (B) qPCR products were analyzed for correct size using 1% agarose gel electrophoresis. (C, D) Effects on cell cycle progression of HEK-KARE1c and HEK293 by nucleic acid staining with propidium iodide followed by flow cytometry analysis. Evaluation of the cell cycle states over the typical production time of 72 h of the cultivation. Percentage of cell in different cell cycle phases are presented for each cell line. HEK293 TF represent triple-transfected HEK293 cells for AAV production. Analysis was performed using ModFit LT.

It is known from literature that gene products of adenoviral elements can contribute to a reduced cell viability and also have an influence on cell cycle progression.[145] It was described that an accumulation in the S-phase occurs upon expression of the ORF6 protein 34k.[146] To fully characterize our new KARE1 cell line a cell cycle analysis was performed. From Figure 40C and D an accumulation in the S-phase was not observed for all clones. The shown distribution into cell cycle states over the typical production time period of 72h. Comparison between the HEK293 cells lines and the KARE1 clones shows no difference. A triple-transfected HEK293 sample was used as control, because adenoviral elements should be expressed here. Straight after transfection a high accumulation in G1 phase was observed. After longer incubation, the cells adapt to the new conditions and show a similar behavior compared to the other samples.

After complete characterization of the new cell HEK293-KARE1 cell line the AAV production ability was assessed. First small-scale transfections were carried out to optimize the production process and we were able to reduce to the total DNA amount and the contamination by the antibiotic blasticidin. Finally, we produced rAAV2 harboring the fluorescence reporter mVenus under control of a CMV promoter as genomic payload in larger scales and purified the preparation using the

established discontinuous iodixanol ultracentrifugation protocol. The final product was analyzed for genomic titer and transduction ability. Genomic titers obtained seem to be low in comparison to production yields using the three-plasmid system (Figure 41) but this can be related to a small number of 100 mm dishes that were used for transfection. The transducing titer is as expected one potency below the genomic titer and indicates that rAAV2 vectors produced using HEK293-KARE1 have the same properties regarding transduction ability. The efficiency of transduction was determined using flow cytometry analysis where the fluorescence of the gene of interest mVenus was measured. Different dilutions of the rAAV2 preparation were applied to HT1080 cells and incubated before detection of the fluorescence signal.

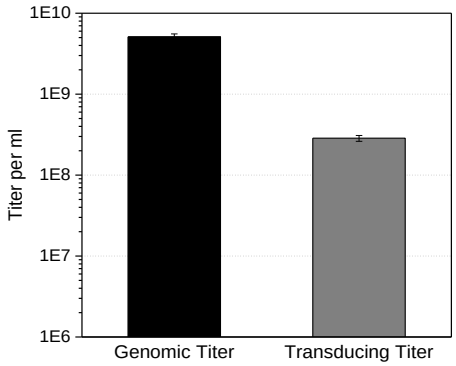


Figure 41: Comparison of genomic and transducing titer for rAAV2 produced using the new HEK-293-KARE cell line. Genomic titers were determined in qPCR reaction. Transducing titers were calculated from transduction efficiencies determined after transduction of HT1080 cells. Transduction efficiency was measured upon expression of the delivers transgene mVenus using flow cytometry analysis.

In summary, a new AAV production cell line was generated in collaboration with Kathrin Teschner and Julian Teschner. A detailed characterization of this cell line was performed before analysis of its production capability. A further collaboration with Axel Rossi from the Büning group in Hanover showed that further optimization needs to be carried out to ensure reproducibility of the protocols. To this end, he provided RepCap and ITR plasmid from their laboratory and we analyzed it in context of our cell line showing that rAAV production is possible.

Finally, we yielded a cell line that has a great potential for simplification of rAAV production because it reduces the amount of DNA used for transfection. There are also further applications where this cell line offers new possibilities, e.g. directed evolution approaches using rAAVs.[147,148] Previous works rely on either co-transduction using the adenovirus, which is considered to be biosafety level 2 or requires the transfection of cells with pHelper plasmid, which results in transfection stress for the cells. Furthermore, adaptation of this cell line to suspension might offer more possibilities in scale-up potential.

6. Conclusion and Outlook

Recombinant adeno-associated viruses are a promising platform for gene therapy approaches. Gaining a deeper insight into the mechanism of re-targeting via rationally designed peptide motifs towards the EGF receptor was the main aim during this work.

The continuation of this project required in first steps an optimization of the AAV production and purification protocols used so far. Standard methods, e.g. ultracentrifugation, have been established which allow a higher quality of the produced AAV preparations. In this context, plasmid systems providing the opportunity to produce AAV2, AAV6 and AAV9 with capsid modifications have been developed and furthermore the production in the suspension cell line 293F was analyzed. The plasmid systems for AAV6 and AAV9 still allow for a broad range of modification that could be analyzed in the future, e.g. presentation of larger proteins as DARPins.

In addition, the EGF receptor density on the available cell lines was determined after the development and characterization of an EGF-mCherry fusion protein. During the course of cell line development, e.g. A431-AAVRko, changes in protein expression could occur and thus influence receptor density. Especially in case of cell lines used for EGFR transduction experiments a change in receptor density needs to be prevented.

Analysis of rationally designed peptides was the focus of this work and started with analysis of the synthetic peptides. Cyclization of the peptides occurred via a disulfide bridge and further attachment of a carboxyfluorescein-label allows for direct visualization in cell culture experiments. All characterization experiment showed that the β-hairpin pep1osy is the most promising rationally designed peptide for EGFR targeting. Here the well-known problem of random diffusion into the cell was observed in flow cytometry and confocal microscopy experiments.

A megadalton complex such as the rAAV inhibits random diffusion and thus provides an optimal scaffold the peptide. Furthermore, decoration of the rAAV capsid with 60 peptides increases the avidity towards the targeted receptor. Integration of the sequences on the genetic level allowed for production of rAAV peptide variants. Both rational designed peptides were able to target the particles towards EGFR-overexpressing cells, while cells with low presentation level of EGFR were not reached. This strategy was also transferred to AAV6 and AAV9 but showed contradictory results. A more detailed characterization of integration sites in AAV6 and AAV9 is necessary to prevent those problems in re-targeting strategies. Finally, the rAAV peptide variants were tested in *in ovo* CAM assays and for AAV2 pep1jhf an increased delivery towards the tumor tissues was observed. Various side effects, e.g. cerebral haemorrhages, have been observed during the experiments. To prevent or reduce them, a detailed analysis of subsequent experiments should be performed. Findings might help to design better *in ovo* working variants in the future. A combination with transcriptional and translational approaches might also increase the specificity towards the tumor tissue.

7. Experimental part

All reaction conditions, experimental procedures and analytical results not listed in this chapter can be found in the publications.

7.1. Peptide synthesis

The resin was deprotected twice with 20% piperidine/DMF (*N*,*N*-dimethylformamide), 0.1 M 1-hydroxybenzotriazole (HOBt) for 20 min and washed. Fmoc-Cys(Trt)-OH (4 equiv) and ethyl cyano-hydroxyiminoacetate (Oxyma) (4 equiv) were dissolved in DMF, combined, *N*,*N*′-diisopropylcarbodiimide (DIC) (4 equiv) was added and the mixture was shaken for 0.5 min. This solution was added to the resin and incubated for 2 h and rt. Afterwards the resin was filtered, washed and capping was carried out twice with acetic anhydride (Ac_2O) (10 equiv) and pyridine (10 equiv) in DMF.

7.1.1. Loading of a Rinkamide resin

Rink-amide resin (0.5 mmol/g) was swollen in DMF (10 ml/g) for 15 min. Prior to coupling the first amino acid, the resin is incubated in 20% piperidine in DMF, 0.1 m 1-hydroxybenzotriazole (HOBt) for 2× 20 min to remove the terminal Fmoc protecting group. The respective amino acid (4 equiv) and Oxyma (4 equiv) were dissolved in DMF, combined, DIC (4 equiv) was added and the mixture was shaken for 0.5 min. This solution was added to the resin and incubated for 2 h at rt. Afterwards the resin was filtered, washed and capping was carried out twice with Ac2O (10 equiv) and pyridine (10 equiv) in DMF. To determine the loading capacity of the resin, 1 mg of the dried resin is mixed with 3 ml 20% piperidine in DMF in a 1 cm quartz cuvette and the absorption of the reaction solution is determined at 290 nm (A290 nm) after 20 min of incubation. As reference for measurements 20% piperidine in DMF was used. The loading capacity was calculated as described in this equation: $\text{Loading capacity}\left[\frac{mmol}{g}\right] = \frac{A_{290\,nm}}{m_{resin[mg]} \times 1.65}$.

7.1.2. Synthesis of triazolyl-bridged peptides

7.1.2.1. Microwave-assisted solid phase peptide synthesis

The further synthesis was conducted automatically using the microwave-assisted peptide synthesizer Liberty (CEM Corporation). All amino acids were double coupled using TBTU (*O*-(benzotriazol-1-yl)-*N*,*N*,*N*',*N*'-tetramethyluronium tetrafluoroborate) and DIPEA (*N*,*N*-diisopropylethylamine) with *N*-terminal acetylation using a mixture of acetic anhydride (Ac_2O), DIPEA and HOBt.

7.1.2.2. *N*-terminal 5(6)-carboxyfluorescein coupling

Peptides were labelled after cyclization on-resin with 5(6)-carboxyfluorescein. Oxyma (3 equiv) and 5(6)-carboxyfluorescein (3 equiv) were solubilized in DMF before addition of DIC (3 equiv). The resin was equilibrated in DMF before addition of the solution. The reaction mixture was stirred

at RT for 12 h protected from light. The resin was washed with DMF (5 x), 20% piperidin in DMF (1 x, 2 min incubation), 20% piperidin in DMF (1 x, 10 min incubation) and DMF (5 x).

7.1.2.3. Cleavage of the peptide from the solid support and final deprotection

After completion of the synthesis, the peptide was cleaved off the resin with a solution of 95% trifluoroacetic acid (TFA), 2.5% water and 2.5% triisopropylsilane (TIPS) in the presence of a spatula tip of dithiothreitol (DTT) at room temperature for 5 h. Full vacuum was applied to the crude material before addition of a 1:1 (*v/v*) acetonitrile and water mixture. The crude peptide was purified by C18 RP-HPLC with a linear acetonitrile gradient (5-100% in 55 min, 0.1% TFA). Finally, the TFA salt of the peptides was isolated.

7.1.3. Synthesis of disulfide-bridged peptides

7.1.3.1. Manual solid phase peptide synthesis

Peptide synthesis was conducted according to standard protocols using the Fmoc/^{t}Bu-strategy in a plastic syringe fitted with a polypropylene porous disk at rt. Washing steps were performed after each reaction with DMF (5×) with 10 ml/g resin. Solvents and soluble reagents were removed by suction. During couplings the resin was shaken with a horizontal shaker. SPPS was conducted with scales at 0.07 mmol using 10 ml/g resin for reactions. Reaction control was performed by MALDI or LC-MS and Kaiser test. Peptides containing 5(6)-carboxyfluorescein were handled in the dark. Rink-amide resin (0.5 mmol/g) was swollen in DMF (10 ml/g) for 15 min. For Fmoc-deprotection the resin was treated twice with a solution of 20% piperidine/DMF, 0.1 M HOBt for 20 min. Coupling steps were repeated twice. Fmoc-Xaa-OH (4 equiv) and Oxyma (4 equiv) were dissolved in DMF, combined and DIC (4 equiv) were added. After inverting the solution for 0.5 min, it was incubated with the resin for 60 min.

Coupling of 5(6)-carboxyfluorescein (3 equiv) with Oxyma (3 equiv) and DIC (4 equiv) was carried out overnight in the dark. Afterwards the resin was treated twice with a solution of 20% piperidine/DMF for 20 min. For peptide pep1osy coupling was incomplete and repeated with 12 equiv of 5(6)-carboxyfluorescein and PyAOP ((7-azabenzotriazol-1-yloxy)tripyrrolidinophosphonium hexafluorophosphate) in a mixture of DMF and NMP and incubated in the dark for 72 h.[149]

7.1.3.2. Cleavage from the resin

Peptides were cleaved from the resin with a mixture of TFA/H_2O/TIPS/DTT (92.5:2.5:2.5:2.5 *v/v/v/w*) for 2×3 h each. For EDA the cleavage solution was purged with argon. The solvent was evaporated under reduced pressure, peptides 1 and 2 were additionally precipitated in ice cold ether, freeze dried and linear peptides are purified by RP-HPLC.

7.1.3.3. Cyclization of peptides

Linear peptides pep1osy and pep1jhf were diluted at 1 mg/ml with 0.1 M $NaHCO_3$ buffer (pH 7.8) and stirred open to atmosphere until completion (2-24 h). Cyclic peptides were desalted using a manual C_{18} column (25 g, 400-220 mesh) using water/0.1% TFA for desalting (200 ml) and MeOH/0.1% TFA (300 ml) for eluting. The solvent was evaporated under reduced pressure and the

crude peptides were freeze dried before purification using RP-HPLC. For pepEDA, the linear peptide was dissolved in degassed MeOH/water (5:95) at 1 mg/ml. A solution of I_2 in MeOH (0.06 M) was added dropwise until the solution stays slightly yellow (~ 3 ml).[150] After completion (0.5 to 1 h) the reaction was quenched with 1 M ascorbic acid obtaining a colorless solution and the solvent was removed under reduced pressure. Crude peptides were freeze-dried and purified by RP-HPLC. For fluorescent labeled peptides experiments and purification were performed in the dark.

7.1.4. General protocols for peptide analysis

Analytical RP-HPLC was performed on a Shimadzu Nexera XR UHPLC equipped with a pump LC-20AD, an autosampler SIL-20AXR, a column oven CTO-2CA, a diode array detector SPD-M20A and a communication module CBM-20A using a Phenomenex Luna C_{18} column (3.0 µm, 100×2.0 mm). Analytical LC-MS (Phenomenex Luna C_{18} column, 3.0 µm, 100×2.0 mm) and determination of HRMS (C_{18} Hypersil Gold column; 1.9 µm, 50×2.1 mm) was performed on an Agilent 6220 TOF-MS with a dual ESI-source, 1200 HPLC system with autosampler, degasser, binary pump, column oven, diode array detector operating with a spray voltage of 2.5 kV. Nitrogen generated by a nitrogen generator NGM11 served both as nebulizer gas and dry gas. External calibration was performed with ESI-L Tuning Mix. MALDI-TOF-MS was conducted with an Ultraflex (Bruker), 355 nm Nd:YAG laser, 50 Hz, positive mode, 1000 shots/spectrum using DHB or CHCA as matrix, calibration with PEG 400-1200. Preparative HPLC was performed using a Merck-Hitachi LaChrom HPLC consisting of interface D-7000, pump L-7150, detector L-7420 and a Hypersil Gold C_{18} column (1.9 µm, 250×21.2 mm) or a Hypersil Gold C_{18} column (7 µm 250×10.0 mm).

7.2. Molecular-biological methods

7.2.1. Oligonucleotides

Oligonucleotides used in this work were purchased at Sigma Aldrich. Molecular cloning procedures are described in the manuscripts and in more detail in the following chapter.

	Name	Sequence 5'- 3'
RFE1	Cap2 587 1jhf for	GATCCGTATC TACCAACCTC CAGGCTGGCA ACTGGGCCAA GAGCCAGGGC AACAAGAGCG AGTACCAGGCC CAAGCAGCTA CAG
RFE2	Cap2 587 1jhf rev	CTGTAGCTGC TTGGGCCTGG TACTCGCTCT TGTTGCCCTG GCTCTTGGCC CAGTTGCCAG CCTGGAGGTT GGTAGATACG
RFE3	Cap2 587 1osy for	GATCCGTATC TACCAACCTC CAGGCTGGCA ACGTGGTGTT CGAGGTGAAC GGCAGAGACC TGGGCTGGGC CCAAGCAGCT ACAG
RFE4	Cap2 587 1osy rev	CTGTAGCTGC TTGGGCCCAG CCCAGGTCTC TGCCGTTCAC CTCGAACACC ACGTTGCCAG CCTGGAGGTT GGTAGATACG
RFE5	Cap2 587 EDA for	GATCCGTATC TACCAACCTC CAGGCTCTGT ACAACCCCAC CACCTACCAG ATGGACGCCC AAGCAGCTAC AG
RFE6	Cap2 587 EDA rev	CTGTAGCTGC TTGGGCGTCC ATCTGGTAGG TGGTGGGGT TGTACAGAGC CTGGAGGTTG GTAGATACG
RFE51	NheI-Start-EGF-for	AAAAAGCTAG CATGAACAGC GACAGCGAGT GCCC
RFE52	GGSG_EGF_rev	AGATCCTCCA CCAGATCCAC CACCCCTCAG CTCCCACCAC TTCAGG

RFE53	Linker-mCherry-for	GGTGGTGGAT CTGGTGGAGG ATCTATGGTG TCCAAGGGCG AAGAGG
RFE54	mCherryXhoI-Rev	AAAACTCGAG TCAGTGGTGG TGGTGGTGGT GCTTGTACAG CTCATCCATG CCGC
RFE56	Cap6/9 SwaI for	ATAATAATTT AAATCAGGTA TGGCTGCCG
RFE57	RFC10 rev	ATAATACTGC AGCGGCCGCT ACTA
RFE58	EcoRV MscIdel for	GCGAGCTCGA TATCAAATTA CGCC
RFE59	CAT MscIdel rev	CGTGGCTAAT ATGGACAACT TCTTCG
RFE60	CAT MscIdel for	CGAAGAAGTT GTCCATATTA GCCACG
RFE61	EcoRI MscIdel rev	GCCGCGAATT CCAGAAATCA
RFE62	Cap9 453 His for	CCGGATTAGG CCATCATCAT CATCATCATA GCCAGAATCA ACAAACGCTA AAATTCAGTG TGGCCGGACC CAGCAACATG GCT
RFE63	Cap9 453 His rev	GTACAGCCAT GTTGCTGGGT CCGGCCACAC TGAATTTTAG CGTTTGTTGA TTCTGGCTAT GATGATGATG ATGATGGCCT AAT
RFE64	Cap9 587 His for	CCACAAACCA CCAGAGTGCC CAATTAGGCC ATCATCATCA TCATCATAGC GCACAGGCGC AGACCGGCTG GGTTCAAAACC AAGG
RFE65	Cap9 587 His rev	GATCCCTTGG TTTTGAACCC AGCCGGTCTG CGCCTGTGCG CTATGATGAT GATGATGATG GCCTAATTGG GCACTCTGGT GGTTTGTGG
RFE66	Cap9 587 EDA for	CCACAAACCA CCAGAGTGCC CAACTGTACA ACCCCACCAC CTACCAGATGG AC-GCACAGGC GCAGACCGGC TGGGTTCAAA ACCAAGG
RFE67	Cap9 587 EDA rev	GATCCCTTGG TTTTGAACCC AGCCGGTCTG CGCCTGTGCG TCCATCTGGT AG-GTGGTGGG GTTGTACAGT TGGGCACTCT GGTGGTTTGT GG
RFE68	Cap9 587 1jhf for	CCACAAACCA CCAGAGTGCC CAATGGGCCA AGAGCCAGGG CAACAAGAGC GAG-TACCAGG CACAGGCGCA GACCGGCTGG GTTCAAAACCA AGG
RFE69	Cap9 587 1jhf rev	GATCCCTTGG TTTTGAACCC AGCCGGTCTG CGCCTGTGCCT GGTACTCGCTC TTGTTGCCCT GGCTCTTGGC CCATTGGGCA CTCTGGTGGT TTGTGG
RFE70	Cap9 587 1osy for	CCACAAACCA CCAGAGTGCC CAAGTGGTGT TCGAGGTGAA CGGCAGAGAC CTGGGCTGGG CACAGGCGCA GACCGGCTGG GTTCAAAACC AAGG
RFE71	Cap9 587 1osy rev	GATCCCTTGG TTTTGAACCC AGCCGGTCTG CGCCTGTGCC CAGCCCAGGT CTCTGCCGTT ACCTCGAACAC CACTTGGGCA CTCTGGTGGT TTGTGG
RFE72	CAT BspEI del for	CATACGAAAT TCCGGGTGAG CATTC
RFE73	CAT BspEI del rev	GAATGCTCAC CCGGAATTTC GTATG
RFE78	Cap6 453His for	CCGGATTAGG CCATCATCAT CATCATCATA GCAGTGCCCA AAACAAGGAC TTGCTGTTTA GCCGGGGGTC TCCAGCTGGC ATGTCT
RFE79	Cap6 453His rev	GTACAGACAT GCCAGCTGGA GACCCCCGGC TAAACAGCAA GTCCTTGTTT TGGG-CACTGC TATGATGATG ATGATGATGG CCTAAT
RFE80	Cap6 587His for	CCACCGAAAG ATTTGGGACT GTGGCAGTCA ATCTCCAGAG CAGCAGCTTA GGCCATCATC ATCATCATCA TAGCACG
RFE81	Cap6 587His rev	GATCCGTGCT ATGATGATGA TGATGATGGC CTAAGCTGCT GCTCTGGAGA TTGACTGCCA CAGTCCCAAA TCTTTCGGTG G
RFE82	Cap6 5871jhf for	CCACCGAAAG ATTTGGGACT GTGGCAGTCA ATCTCCAGAG CAGCAGCTGG GCCAAGAGCC AGGGCAACAA GAGCGAGTAC CAGACG
RFE83	Cap6 5871jhf rev	GATCCGTCTG GTACTCGCTC TTGTTGCCCT GGCTCTTGGC CCAGCTGCTG CTCTG-GAGAT TGACTGCCAC AGTCCCAAAT CTTTCGGTGG
RFE84	Cap6 5871osy for	CCACCGAAAG ATTTGGGACT GTGGCAGTCA ATCTCCAGAG CAGCAGCGTG GTGTTCGAGG TGAACGGCAG AGACCTGGGC TGGACG
RFE85	Cap6 5871osy rev	GATCCGTCCA GCCCAGGTCT CTGCCGTTCA CCTCGAACAC CACGCTGCTG CTCTG-GAGAT TGACTGCCAC AGTCCCAAAT CTTTCGGTGG

7.2.2. General cloning procedures

Modifications of the 453 and 587 loop were carried out using the unique restriction enzyme sites for all AAV serotypes. The Rep2Cap2 plasmid was already equipped with unique SspI and SalI (453 region) as well as PvuII and BamHI (587 region). Plasmids for serotype 6 and 9 were created starting from the Rep2Cap2 plasmid. Cap6 and Cap9 were ordered as string syntheses at GeneArt (ThermoFisher Scientific). Before introduction into the pSB1C3_001-Rep2 plasmid, restrictions sites, enabling loop modifications, had to be eliminated from the plasmid backbone. By PCR reaction MscI (RFE60+RFE61) and BspEI (RFE72+RFE73) restriction sites were eliminated from the CAT marker region. The oligonucleotide pair RFE56 and RFE57 was used to amplify Cap6 and Cap9 from strings before ligation with the Rep2 plasmid which was opened with SwaI and PstI. Introduction of peptide motifs in loop regions is possible using the now unique restriction sites BspEI/BsrGI (453 region) and MscI/BamHI (587 region). Complementary oligonucleotides with overlapping ends between those recognition sites were designed, that match the overhangs generated during a restriction digest. After phosphorylation (T4 PNK) and annealing of complementary oligonucleotides a subsequent ligation into the digested and dephosphorylated (AnP) vector can occur.

Table 3: Standard protocols for common cloning procedures. Incubation temperatures or detailed procedures are chosen according to the manufacturer's instructions.

Restriction digest, preparative		**Phusion PCR**	
CutSmart buffer	5 µl	GC buffer	10 µl
Enzyme	1 µl	10 mM dNTPs	1 µl
Plasmid	1 µg	DNA template	100 ng
Water	until 50 µl	10 µM Primer for/rev	each 2.5 µl
		DMSO (final 3%)	0.5 µl
		Phusion HF polymerase	0.5 µl
		Water	to 50 µl
T4 DNA ligase		**pJET blunt protocol**	
Ligase buffer	2 µl	(CloneJet PCR cloning Kit #K1231)	
T4 DNA Ligase (5 U/µl)	1 µl	2× Reaction buffer	5 µl
Vector DNA	100 ng	PCR product	4 µl
Insert DNA	1:1 to 5:1 molar ratio over vector	pJET1.2	0.5 µl
Water	to 20 µl	T4 ligase	0.5µl

Polymerase chain reactions (PCRs) were carried out using Phusion polymerase (NEB) with the standard protocol described by the manufacturer. All restriction and cloning enzymes were purchased at NEB and used according to manufacturer's instructions. For ligation reactions a T4 ligase from ThermoFisher Scientific was applied with the provided manual. Plasmids were transformed into chemically competent *E. coli* DH5α cells, plated on LB agar plates with the corresponding antibiotic and incubated at 37 °C overnight. Plasmid DNA was isolated from bacterial colonies using the NucleoSpin Plasmid Kit from Macherey-Nagel. Cloning success was verified via sequencing at the Bielefeld University Sequencing Core Facility. Larger quantities of plasmid DNA for transient transfections were generated according to the instructions for the NucleoBond Xtra Midi Kit from Macherey-Nagel.

7.2.3. Plasmids

Table 4: List of plasmids used during this work.

Number	Name	Source
pZMB0088	pHelper	Agilent
pZMB0091	pSB1C3_001_pCMV_DARPinE01_mli_VP23 _453_587koHis	iGEM
pZMB0154	pSB1C3_001_Rep_VP123_453_587ko_p5tataless	iGEM
pZMB0155	pSB1C3_001_Rep_VP13_453_587ko_p5tataless	iGEM
pZMB0165	pSB1C3_001_Rep_VP123_453_587ko1jhf_p5tataless	RFE master thesis
pZMB0166	pSB1C3_001_Rep_VP123_453_587ko1osy_p5tataless	RFE master thesis
pZMB0193	pSB1C3_001_Rep_VP123_453_587koEDA_p5tataless	RFE master thesis
pZMB0246	pSB1C3_001_CMV_VP1up_NLS_mVenus_VP23_453_587koHis	iGEM
pZMB0307	pSB1C3_001_pCMV_Kozak_VP23_453_587wt	KSC
pZMB0490	pET21a-EGF-mCherry-His6	This work
pZMB0493	pSB1C3_001_FUCA1	This work
pZMB0494	pSB1C3_001_FUCA1_T2A_EGFP	This work
pZMB0495	pSB1C3_FUCA1_T2A_EGFP_hGHpA	This work
pZMB0496	pSB1C3_CMV_FUCA1_T2A_EGFP_hGHpA	This work
pZMB0497	pUC19bb_ITR_EXS_CMV_FUCA1_T2A_EGFP_hGHpA	This work
pZMB0522	pUC19bb_ITR_EXS_pCMV_mVenus_hGHpolyA	PBO
pZMB0550	pSB1C3_002_RepCap2_VP123_453_587wt_p5tataless	This work
pZMB0551	pSB1C3_002_RepCap9_VP123_453_587_p5	This work
pZMB0552	pSB1C3_002_RepCap9_VP123_453His_587_p5	This work
pZMB0553	pSB1C3_002_RepCap9_VP123_453_587His_p5	This work
pZMB0554	pSB1C3_002_RepCap9_VP123_453_587EDA_p5	This work
pZMB0555	pSB1C3_002_RepCap9_VP123_453_5871jhf_p5	This work
pZMB0556	pSB1C3_002_RepCap9_VP123_453_5871osy_p5	This work
pZMB0573	pSB1C3_002_RepCap6_VP123_453_587wt	This work
pZMB0574	pSB1C3_002_RepCap6_VP123_453_587wt_p5	This work
pZMB0575	pSB1C3_002_RepCap6_VP123_453_587wt_p5tataless	This work
pZMB0576	pSB1C3_002_Rep(PVC)Cap9_VP123_453_587_p5	This work
pZMB0592	pSB1C3_002_Rep2Cap6_453His_587_p5	This work
pZMB0593	pSB1C3_002_Rep2Cap6_453_587His_p5	This work
pZMB0594	pSB1C3_002_Rep2Cap6_453_5871jhf_p5	This work
pZMB0595	pSB1C3_002_Rep2Cap6_453_5871osy_p5	This work

7.3. Biochemical methods

7.3.1. Recombinant protein expression in *E. coli* and 293F

During this work, recombinant proteins were expressed using two different expression systems depending on the protein to be expressed.

The fusion protein EGF-mCherry was expressed using the prokaryotic pET system in *E. coli* BL21(DE3). The plasmid was transformed into chemical competent *E. coli* and a preculture was grown overnight (37 °C, 180 rpm) before inoculation of 0.5 l LB medium containing 100 µg/ml ampicillin. This culture was grown (37 °C, 180 rpm) up to an OD_{600} of 0.5 before induction of

protein expression with isopropyl-β-D-thiogalactopyranoside (IPTG) (0.1 mM). Cells were cultivated at 30 °C, 180 rpm for 4 h. Afterwards, cells were harvested by centrifugation (3220×*g*, 30 min), washed with 100 mM Na_2HPO_4 buffer (pH 7.4) and stored at -20 °C.

The soluble EGFR was expressed from eukaryotic cells using suspension adapted 293F cells (Freestyle 293-F, Thermo Fisher Scientific) in chemically defined animal component free medium (Xell AG). Prior to transfection of 293F cells, cells were washed with PBS. Transient transfection of 3×10^6 cells/ml was carried out with polyethylenimine Max in a 1:4 DNA-PEI ratio (w/w) with 2 µg DNA (0.6 pg per cell). Transfections were carried out in 125 ml shake flasks in a volume of 30 ml. The cultivation of cells lasted until the viability of cells dropped below 60%. Cells were harvested at 2000×g for 5 min, the cell culture supernatant was sterile filtered and applied to subsequent purification as recombinant proteins are secreted from the cells. The supernatant was stored at -20 °C until purification.

7.3.2. Purification of recombinant protein

Recombinant proteins used during this work were expressed with a terminal His_6-tag allowing the purification using immobilized metal affinity chromatography (IMAC) on a Ni-NTA resin (Macherey-Nagel).

For protein purification, cells from *E. coli* cultivation were thawed and suspended in 30 ml equilibration buffer (50 mM Na_2HPO_4, 300 mM NaCl, 10 mM imidazole pH 8). The cell membrane of bacteria was disrupted in three repeated French press cycles at a pressure of 1000 psi. Cell debris was separated from soluble protein by centrifugation at 15.000×g for 30 min at 4 °C and the supernatant filtered subsequently.

The cell lysate after French press and the cell culture supernatant from 293F expression can both be applied to subsequent chromatography. The solution was applied to a self-packed Ni-NTA column using an ÄKTA start protein purification system. After washing with 30 CV buffer containing 50 mM imidazole in equilibration buffer, the protein was eluted with an elution buffer containing 300 mM imidazole. If not further purified using other methods, the eluted protein was concentrated in PBS using Amicon Ultra centrifugal filter units (Merck Millipore) (MWCO corresponding to the molecular weight of the recombinant protein).

If a higher purity was required or the first step purification did not result in satisfying purity an ion exchange chromatography (IEX) was performed on an ÄKTA start protein purification system. The protein was applied to a self-packed Q Sepharose Fast Flow column (GE Healthcare) with a column volume of 1 ml and a gradient from 0 to 2000 mM NaCl in 50 mM Tris. The pH value of the Tris buffer was chosen dependently on the recombinant proteins characteristic isoelectric point (pI). Fractions with a volume of 1 ml were collected. Finally, eluted protein was concentrated in PBS using an Amicon Ultra centrifugal filter unit.

7.3.3. SDS-PAGE and Western blot

Polyacrylamide gels for SDS-PAGE were generated using the Hoefer SE260 system. Samples with 5× SDS loading buffer (20% (*w/v*) glycerin, 200 mM DDT, 10% (*v/v*) SDS, 0.2 M Tris-HCl, 0.05% (*w/v*) bromophenol blue) were prepared and incubated at 95 °C for 10 min followed by centrifugation for 1 min at 10,000×g and finally 20 µl per lane were loaded on a SDS-polyacrylamide gel. Electrophoresis conditions were 100 V for 30 min before increasing to 180 V for 60 min according to the percentage of the gel. Polyacrylamide gels were stained with Coomassie Blue staining solution if not used for western blot analysis.

Samples were blotted onto a 0.45 µm nitrocellulose membrane (Thermo Fisher Scientific) using semi-dry electrophoretic transfer (V20-SDB, Sci Plas). Transfer conditions were chosen according to the size of the protein and the SDS-PAGE gel. For a standard gel (10%) and protein in the range of 30-80 kDa 4 mA/cm^2 were applied for 30 min. After blocking the membrane with 10% (*w/v*) non-fat milk in TBS, the membrane was incubated for 1.5 h with a primary antibody in blocking buffer (see antibodies and dilutions in Table 5). After subsequent incubation with a secondary anti-mouse IgG linked to a horseradish peroxidase, blots were imaged by luminescence detection (Pierce ECL Western Blot Substrate, Thermo Fisher Scientific).

Table 5: Primary antibodies used for Western Blot analysis.

Antibody	Host	Dilution	Supplier
anti-AAV VP1/VP2/VP3, B1, supernatant, monoclonal	Mouse	1:100	Progen
anti-AAV2 (intact particle), A20R, recombinant	Mouse	1:250	Progen
β-Actin (8H10D10), monoclonal	Mouse	1:1000	Cell Signaling Technology
Tetra·His antibody	Mouse	1:2000	QIAGEN
Anti-mouse IgG, HRP-linked Antibody	Horse	1:5000	Cell Signaling Technology

7.3.4. mRNA quantification using qPCR

Total cellular RNA from HEK293 and HEK293-KARE1 clones was isolated using Trizol Reagent (Invitrogen). 5×10^6 cells were pelleted and 1 ml Trizol (ThermoFisher Scientific) was added. The pellet was resuspended, vortexed for 35 s and incubated at room temperature for 5 min before another 15 s of vortexing. Samples were mixed for 15 s after the addition of 200 µl chloroform. A 3 min incubation was followed by centrifugation at 12,000×g for 15 min at 4 °C. The aqueous phase (approx. 500 µl) was separated in a new tube and an equal amount of 100% *iso*-propanol was added. The sample was incubated at room temperature for 10 min before the top phase is removed and 1 ml of 75% ethanol was added. After strong vortexing, the sample was centrifuged at 7,500×g for 5 min at 4 °C. The supernatant was aspirated and the pellet was dried for 5-10 min before resuspending in 35 µl RNase-free water. Incubation at 55 to 60 °C for 10-15 min increased solubility of RNA. 1 µg of RNA was subsequently digested by 1 U DNase I (ThermoFisher Scientific) for 30 min at 37 °C. cDNA was synthesized using oligo(dT)$_{18}$ primer and Revert Aid M-MuLV reverse transcriptase (ThermoFisher Scientific). Quantitative real-time PCR using GoTaq® qPCR Master Mix (Promega) was performed with a LightCycler 480 II detection system (Roche) according to

the manufacturer's protocol. The following primer pairs were used to detect E1A (E1A for 5'-AACCAGTTGC CGTGAGAGTTG-3'; E1A rev 5'-CTCGTTAAGC AAGTCCTCGA TACAT-3'), E2A (E2A for 5'-TCAGGCACAA CCATCCGCGG-3'; E2A rev 5'-GGTCGGGCGC CGA-TATCTTGA-3') and E4 (E4 for 5'-GAACCCTAGT ATTCAACCTG CCACCTCCC-3'; E4 rev 5'-CCACACGGTT TCCTGTCGAGCC-3'). All samples were run in triplicate. C_t values of technical replicates were averaged and averages were used for further calculations. The relative expression $\Delta\Delta C_t$ of E2A and E4 in HEK293-KARE1 cells was calculated from the differences in HEK293 and HEK293-KARE1a to HEK293-KARE1c after normalizing to E1A gene expression $\Delta\Delta C_t = C_t(\text{GeneHEK293-KARE}) - C_t(\text{E1AHEK293-KARE}) - [C_t(\text{GeneHEK293}) - C_t(\text{E1AHEK293})]$.

7.4. Biophysical measurements

7.4.1. Circular dichroism (CD)

CD spectra were recorded with a Jasco J-810 CD-spectrometer equipped with a Peltier-type temperature controller. Measurements of peptides were conducted at a concentration of 100 µM in 10% TFE/water (*v*/*v*). Each spectrum was recorded thrice in the range 190-250 nm at a constant temperature of 25 °C at a scanning rate of 50 nm/min.

Measurements of proteins were conducted in 10 mM Na_2HPO_4, 5 mM NaCl, pH 7.4. Temperature stability measurement were recorded in the range 200 to 240 nm with increasing temperature up to 90 °C. Three spectra were recorded every 1 K with a scanning speed of 50 nm/min and a heating rate of 1 K/min. Data was analyzed using Origin2019 (OriginLab).

7.4.1. Transmission electron microscopy (TEM)

Carbon-coated copper grids, 200 mesh (Electron Microscopy Science) were treated with oxygen plasma (Zepto, Diener electronic GmbH). After this, 3 µl of precipitation-purified rAAV sample[151] was applied to the grid and incubated for 2 min. Excess liquid was drained off, the grid was dried at room temperature and washed with three drops of distilled water. Negative staining was performed using 3 µl 2% (*v*/*v*) uranyl acetate (Science Services) for 30 s. Excess liquid was drained off and grids were dried before channeling the sample into the microscope. rAAVs were visualized with a Philips CM100 (PW6021) instrument with an acceleration voltage of 80 kV. Images were analyzed using the Soft Imaging Viewer (Olympus) and ImageJ.

7.4.2. Atomic force microscopy (AFM)

AFM measurement were performed on a Multimode 8 AFM (Bruker) with Tap300AI-G cantilevers (BudgetSensors) in tapping mode in air. 2 µl of UC purified sample were spotted onto freshly cleaved mice for 1 min before it was rinsed three times with distilled water and dried under a gentle flow of nitrogen. Data analysis was performed with Gwyddion 2.48. Obtained images were treated with offset and plane correction algorithms and the size of visualized particles was measured at half maximum particle height.

7.4.3. Fluorescence polarization (FP)

Fluorescence polarization assays were performed to evaluate the binding affinity of the carboxyfluorescein-labelled peptide (FAM-labelled peptide) towards the soluble EGFR (sEGFR). Series of dilutions were prepared in FP buffer (0.05% Tween®20 in 1× PBS (137 mM NaCl, 2.6 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, pH 7.2) for sEGFR starting with a concentration of 25 µM (2.5× dilution). A volume of 15 µl of the dilutions was pipetted into 384-well plates (BRANDplates®, 384-well, pureGrade™, black). The fluorescent-labelled peptide (20 µM stock) was diluted to a final concentration of 40 nM in FP buffer. A volume of 5 µl peptide dilution (10 nM final) was added to each well. The plate was incubated at 4 °C for 1h before centrifugation at 1000×g, 2 min. Fluorescence polarization was measured using a Tecan Spark 10M with Ex 485/20 nm and Em 520/20 nm. Each sample was measured as a technical triplicate. Polarization was converted into anisotropy and analysis of experimental data was performed in Origin2019 for the normalized values. The 1:1 binding model equation is shown below with r as anisotropy, r_0 as anisotropy of the free peptide, r_b as anisotropy of the EGFR:FAM-labelled peptide, K_d as dissociation constant, $[L]_t$ as total labelled ligand concentration and $[P]_t$ as total protein concentration.

$$r = r_0 + (r_b - r_0)\frac{K_d + [L]_t + [P]_t - \sqrt{(K_d + [L]_t + [P]_t)^2 - 4[L]_t[P]_t}}{2[L]_t}$$

7.4.4. Biolayer interferometry (BLI)

Protein-protein interactions were analyzed using a BLItz system (ForteBio). An ARG2 sensor was equilibrated in water for 10 min. The sEGFR was immobilized on the sensor using EDC/NHS coupling chemistry. The following five steps were performed to immobilize the sEGFR on the sensor: Initial baseline in water (30 s), activation of the surface with 10 µM EDC (1-ethyl-3-(3-dimethylaminopropyl)carbodiimid) and 5 µM NHS (*N*-hydroxysuccinimid) (300 s), loading of sEGFR (1 µg/ml in 10 mM sodium acetate, pH 3.5) (600 s), quenching with 1 M ethanolamine (300 s), regenerating with 50 mM NaOH and final equilibration with water. A loaded sensor can be used for multiple interaction measurements. The scheme for measuring protein-protein interactions is described in the following: Initial baseline in water (30 s), association with varying analyte concentrations (120 s), dissociation in water (180 s), regeneration with 50 mM NaOH (40 s) and final baseline in water (120 s). The EGF-mCherry concentration of the initial solution was determined in three replicates using a NanoDrop2000c (ThermoFisher Scientific). Concentrations of measured samples were estimated from the dilution series. As a negative control for sEGFR binding, bovine serum albumin (BSA) was tested, proving no association. Finally, sensorgrams were fitted globally to a 1:1 binding model by Blitz Pro 1.2.1.3. Curves and global fits were plotted using Origin2019 (OriginLab).

7.5. Cell culture techniques

7.5.1. Cultivation of eukaryotic cells

All cells were maintained at 37 °C with 5% CO_2 and a humidity of 95% at subconfluency in their appropriate medium. Adherent cells were cultivated in T75 flask (Sarstedt) before usage in final experiments. Suspension culture was maintained in tube spin bioreactors (TPP) with orbital shaking at 185 rpm and 5 cm amplitude.

7.5.2. Cryopreservation of eukaryotic cells

Cultures of mammalian cells can be stored long-term in an ultra-low freezer. In a first step the number of viable cells was determined from a subconfluent culture. Cells were pelleted and resuspended in freezing medium to a concentration of 1×10^7 cells/ml. Freezing medium consists of the complete medium containing 10% dimethylsulfoxide (DMSO). Cells were aliquoted to 1 ml in cryogenic vials and frozen slowly at 1 °C/min in an insulated freezing container in a -80 °C freezer. For long-term storage vials were transferred to an ultra-low freezer (-150 °C)

Cryopreserved cells were thawed immediately in a 37 °C water bath. The cells were added to 20 ml of PBS and centrifuged for 5 min at 700xg. The supernatant was discarded and 20 ml of fresh medium supplemented with 10% fetal calf serum (FCS) and 1% P/S (P4333, Sigma Aldrich).

7.5.3. Counting and seeding of eukaryotic cells

The number and viability of mammalian cells was determined using the Luna Automated Cell Counter system (Logos Biosystems). A sample of cells was taken during passaging of cells and mixed with an equal volume of 0.4% trypan blue solution (Sigma Aldrich). A volume of 10 µl was applied to an associated cell counting slide and inserted into the counting system. The device analyses the sample automatically and gives the total cell number, viable cell number and percentage of viability. Results were used to calculate the cell number needed for experiments and finally, a desired cell number was seeded in an appropriate plate.

7.5.4. Confocal microscopy

Live cell imaging was performed on an inverted confocal laser scanning microscope Zeiss LSM780 to visualize binding and internalization of carboxyfluorescein-labelled peptides. Cells ($1.5x10^4$) were seeded in an 8-well µ-slide (Nunc, LabTek) in 300 µl medium and incubated overnight. Microscopy was performed using an incubation chamber at 37 °C to maintain cell culture conditions. Cells were incubated with 10 µl NucBlue Live ReadyProbes Reagent (Invitrogen) and 50 nM Lysotracker DND-99 (Invitrogen) for 10 min. Peptides were diluted to 100 µM stock solution in DMSO and added to the cells with 5 µM concentration for 10 min at 37 °C (final DMSO concentration 5%). Five extensive washing steps with RPMI (w/o phenolred and FCS) were performed before imaging the cells with the 63× objective of the inverted laser scanning Zeiss LSM780 microscope. Laser and detector ranges were used with corresponding main beam splitters as given in

Table 6. Image acquisition was performed with Zeiss Zen 2011 and image analysis with ImageJ. A macro was used to analyze areas of colocalization between the carboxyfluorescein signal of the peptide and the Lysotracker DND-99 signal.

Table 6: Laser and detector ranges used during live cell imaging.

Channel	Color	Laser line	Detector Range
NucBlue (Hoechst 33342)	Blue	405 nm	414-417 nm (Ch1)
Carboxyfluorescein	Green	488 nm (argon)	499-553 nm (ChS1)
Lysotracker-DND99	Red	594 nm	609-700 nm (Ch2)

7.5.5. Wound healing assay

A431 cells (2×10^5 cells/ well) were seeded in 24-well plates (TC-treated, Sarstedt) and cultivated in RPMI-1640 culture media (D8758, Sigma Aldrich; supplemented with 10% FCS and 1% P/S) at 37 °C and 5% CO_2 for 24 h until a confluent monolayer was formed. Cells were serum-starved overnight in RPMI-1640 without FCS. Two scratches per well were made using a sterile 10 µl pipet tip. After washing with PBS thrice, cells were treated with the test substance at a certain concentration (e.g. 1 nM EGF) in RPMI-1640 with 1% P/S. Control cells received an equal volume of buffer (1 ml RPMI-1640 with PBS). Cells were maintained at cultivation conditions and cell migration was observed using a Leica DMI6000 microscope at selected time points ranging from 0 h to 8 h of incubation. The analysis of identical positions is guaranteed by the Mark and Find feature of Leica in combination with a high-precision motorized stage. The area of cell-free surfaces was determined using the MRI Wound Healing Tool plug-in (Volker Bäcker) for ImageJ version 1.51.[152] A linear regression of the surface area versus time was used to determine the cell migration rate in Origin2019 (OriginLab) using $v_{migration} = \frac{|Slope|}{2 \times length\ of\ the\ gap}$. The length of the gap was measured using ImageJ. All experiments were carried out in two independent experiment.

7.6. Virological methods

7.6.1. rAAV vector packaging

rAAV particle were typically produced in adherent HEK293 by transfection using the adenovirus-helper-free plasmid system.[153] The day before transfection, cells were seeded in either 100 mm (3×10^6 cells) or 150 mm (9×10^6 cells) cell culture dishes, respectively. Transfection was performed 24 h after seeding of cells using the calcium phosphate method. Depending of the size of cell culture dish a total DNA amount of 15 µg or 45 µg was used (5 pg DNA per cell). For triple transfection the pHelper, ITR and RepCap plasmid were applied in equimolar concentrations. For one 100 mm dish DNA was added to 0.5 ml of 0.3 M calcium chloride solution and mixed well. The DNA mixture was then added dropwise to sterile 2× HBS buffer (50 mM HEPES, 1.5 mM Na_2HPO_4, 280 mM NaCl, pH 7.05), vortexed thoroughly and immediately added dropwise to the plate of cells. Transfected cells were harvested after 72 h of incubation at 37 °C. Cells and media were separately collected and stored until final purification at -80 °C.

7.6.2. Discontinuous iodixanol density gradient ultracentrifugation

To release viral particles, cell pellets were first resuspended in lysis buffer (50 mM Tris, 150 mM NaCl, 2 mM $MgCl_2$, pH 7.5-8) and then subjected to three freeze-thaw cycles. Unwanted DNA and RNA contamination was digested by benzonase nuclease (50 U per ml lysis buffer) at 37 °C for 30 min following by addition of 0.5% of the detergent CHAPS (3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate hydrate). The samples were centrifuged at 3000×g for 10 min. Viral particles in the spent cell culture medium were isolated from the medium by ammonium sulfate precipitation. Ammonium sulfate (12.52 g per 40 ml volume) was dissolved and incubated for 30 min on ice before centrifugation at 8300×g for 10 min at 4 °C. The pellet after centrifugation was resuspended in the previously generated supernatant from the cell pellet. This final sample was used for iodixanol ultracentrifugation to remove further contaminants. Iodixanol (Progen) solutions of 1.25 ml 60%, 1.25 ml 40%, 1.5 ml 25% and 2.25 ml 15% were layered in an open top polyallomer 16 x 76 mm tube (Science Services). Finally, the sample was applied with a maximal volume of 5 ml. Tubes were sealed with cap assemblies and tared before centrifugation in a T-880 rotor (Sorvall) at 340,000×g for 2 h at 18 °C. The rAAV containing fraction was collected at the interphase between 60% and 40% iodixanol layer with a 21G × 1 1/2" injection needle and buffer was exchanged to 1× HBSS (Sigma Aldrich) via Amicon Ultra-4 100K centrifugal filter units (Merck pore). Aliquots of the final sample were stored at -80 °C before final application in experiments.

Table 7: Composition of iodixanol solution for application in density gradient ultracentrifugation.

Solution	60% Iodixanol	PBS-MK	Additionally
54%	8 ml	0.887 ml 10× PBS-MK	
40%	4 ml	1.4 ml 1× PBS-MK	
25%	2.5 ml	2.9 ml 1× PBS-MK	50 µl Phenolred
15%	1.5 ml	1.2 ml 1× PBS-MK	2.7 ml 1× PBS-MK 2 M NaCl

7.6.3. Determination of genomic titers

Genomic titers of viral vector preparations were determined by qPCR of freshly frozen samples. Prior to qPCR, samples were treated with 10 U DNaseI (New England Biolabs) in 10× DNaseI buffer in a final volume of 50 µl at 37 °C for 30 min before heat inactivation (75 °C, 20 min). For samples from crude lysate additionally 0.8 U Proteinase K (New England Biolabs) was incubated for 50 min at 37 °C before heat inactivation (95 °C, 10 min). Different dilutions of the DNaseI digest were used as template in the qPCR reaction. The sample was mixed with 2.5 µl primer qPCR-hGH-for (5'-CTCCCCAGTG CCTCTCCT-3') and 2.5 µl primer qPCR-hGH-rev (5'-ACTT-GCCCCT TGCTCCATAC-3'), each at a stock concentration of 4 µM, and 10 µl of 2× GoTaq qPCR Mastermix (Promega). The PCR reaction was carried out as described in the manual (TM318 6/14, Promega) with an increased time interval for the first denaturation step (95 °C, 10 min) using a LightCycler 480 II (Roche). The genomic titer was calculated from a standard curve of 10 to 10^6 copies of the ITR plasmid (pZMB0522 containing the hGH polyA motif) with an efficiency between 90-110% and an R value less than 0.1. Genomic titers for crude lysates were estimated from a standard curve spiked with the same amount of a non-transfected cell lysates.

7.6.4. Cell transduction assay

Transduction efficiencies of viral preparations were assayed on cells. Cells to be transfected were seeded in 12-well plates in 500 µl of complete media. Viral samples were applied in different multiplicities of infection (MOI) after cells have settled for about 1 h. On average, MOIs ranging from 100 to 100,000 were used in all experiments. The exact MOI is given in each individual experiment. After overnight incubation, 500 µl of fresh complete medium was added to each well. Cells are incubated for at least 48 h at 37 °C prior to detection of the fluorescence reporter mVenus. Incubation time was increased for modified rAAVs up to 96 h. Fluorescence was detected using fluorescence microscopy with a Leica DMI6000. Afterwards cells were trypsinized and applied to subsequent flow cytometry analysis using a FACScalibur. Here, at least 10,000 events were measured and analyzed using FlowJo V10 software.

7.6.5. *In ovo* chorioallantoic membrane (CAM) assays

Besides production and purification of viral vectors, all steps of the CAM assays have been conducted by Dr. Lea Krutzke, Ulm.

7.6.5.1. CAM assay

Fertilized eggs at embryonic day 0 were carefully cleaned with tap water before incubation at horizontal position at 37.8 °C and 58-60% relative humidity (ProCon Grumbach, Compact S84 with automatic turning trays). Automatic turning of the egg five times by 180° at 6 °C/s prevents vibration. After 96 hours, a small hole was punctured into the pointed end of the egg. The eggs were then incubated for 10 min in an upright position to allow air to escape and deepening of the egg content. Before enlargement of the hole with scissors to a diameter of 1 cm, the hole was sealed with a patch (Leukosilk S). After closing the larger hole with a second patch, eggs were incubated in upright position without turning until day 7 where a 1 mm thick silicon ring with an inner diameter of 5 mm was placed on the chorioallantoic membrane atop a branching vessel. During all incubation steps the opening of the egg is closed with a patch. The day after, a suspension of 3×10^6 A431 cells in a final volume of 15 µl was mixed with 10 µl ice-cold matrigel (Corning) and added to the silicon ring on the CAM. A431 cells were detached from adherent culture using accutase before washing with PBS and resuspension in serum-free medium. At day 12 a second enlargement of the hole took place to allow for intravenous injection of 1×10^{11} rAAVs in 100 µl PBS with a G27x 3/4" / ø 0,40 x 20 mm cannula. After injection the cannula was left at the site of injection to prevent severe bleedings due to poor coagulation of the chick blood. The syringe was removed, the cannula sealed with a silicon plug and fixed with several patches. At day 14 chick embryos were sacrificed by systemic injection of propofol (3 mg per egg) before extraction of tumors and organs. All biologic samples were deep-frozen in liquid nitrogen and stored at -80 °C.

7.6.5.2. DNA isolation from organs

DNA isolation from organs was performed using the GenElute™ Mammalian Genomic DNA Miniprep kit. Organs were thawed and cut into small pieces of about 10 mg before addition of 180 µl

lysis buffer T. Proteinase K (20 µl) is added and the samples are incubated at 56 °C overnight to assure complete digestion of the tissue. Further steps were performed according to the manufacturer's instructions. Finally, DNA is eluted in 200 µl 10 mM Tris, pH 8.0 and stored either at 4 °C (short term) or -80 °C (long term).

7.6.5.3. Quantification of viral DNA by qPCR

Viral DNA is quantified from samples of isolated DNA. The DNA concentration was determined photometrical and samples diluted in ddH_2O to a final concentration of 10 ng/µl. The qPCR reaction was performed using 10 µl Kapa SYBR FAST (# KK4602 Kapa Biosystems), 0.4 µl primer forward (concentration of 10 pmol/µl), 0.4 µl primer reverse and 7.2 µl ddH_2O with 2 µl DNA sample per well (20 ng DNA total). Sequences of oligonucleotide primers for qPCR are given in Table 8. PCR reactions were performed as described by the manufacturer with an increased time interval (10 min) for the initial denaturation step. A standard curve from 1×10^9 to 1×10^2 copy numbers per well was prepared from pZMB0522 and used to finally quantify the viral DNA amount.

Table 8: Oligonucleotide primers used for vial DNA quantification in qPCR reaction. The β-actin primer pair for gallus gallus was only used for organs while the β-actin pair for human was used for the tumor tissue. β-actin served as housekeeping gene for normalization of the data set.

	Name	Sequence 5'- 3'
LK46	ß-Aktin (gallus gallus) for	ATTGCCCCACCTGAGCGCAA
LK47	ß-Aktin (gallus gallus) rev	CATCTGCTGGAAGGTGGACA
LK56	ß-Aktin (human) for	GCTCCTCCTGAGCGCAAG
LK57	ß-Aktin (human) rev	CATCTGCTGGAAGGTGGACA
LK50	mVenus for	AAACTGATTTGCACCACCGG
LK51	mVenus rev	GCAAAGCATTGCAGGCCATA

8. References

(1) WHO. Cancer: Key facts https://www.who.int/news-room/fact-sheets/detail/cancer (accessed Feb 5, 2019).

(2) Ferlay, J.; Colombet, M.; Soerjomataram, I.; Dyba, T.; Randi, G.; Bettio, M.; Gavin, A.; Visser, O.; Bray, F. Cancer Incidence and Mortality Patterns in Europe: Estimates for 40 Countries and 25 Major Cancers in 2018. *Eur. J. Cancer* **2018**, *103*, 356–387.

(3) Ginn, S. L.; Amaya, A. K.; Alexander, I. E.; Edelstein, M.; Abedi, M. R. Gene Therapy Clinical Trials Worldwide to 2017: An Update. *J. Gene Med.* **2018**, *20* (5), 1–16.

(4) FDA Center for Biologics Evaluation and Research. Long Term Follow-Up After Administration of Human Gene Therapy Products. **2018**.

(5) Emdad, L.; Sarkar, D.; Fisher, P. B. Gene Therapies for Cancer: Strategies, Challenges and Successes. *HHS Public Access* **2016**, *230* (2), 259–271.

(6) Amer, M. H. Gene Therapy for Cancer: Present Status and Future Perspective. *Mol. Cell. Ther.* **2014**, *2* (1), 27.

(7) Miller, A. D. Cationic Liposomes for Gene Therapy. *Angew. Angew. Chemie Int. Ed.* **1998**, *37* (13–14), 1768–1785.

(8) Li, S. Electroporation Gene Therapy: New Developments in Vivo and in Vitro. *Curr. Gene Ther.* **2004**, *4* (3), 309–316.

(9) Kotterman, M. A.; Schaffer, D. V. Engineering Adeno-Associated Viruses for Clinical Gene Therapy. *Nat. Rev. Genet.* **2014**, *15* (7), 445–451.

(10) FDA Advisory Committee. *Spark Therapeutics, Inc LUXTURNA (Voretigene Neparvovec)*; 2017.

(11) Srivastava, A.; Lusby, E. W.; Berns, K. I. Nucleotide Sequence and Organization of the Adeno-Associated Virus 2 Genome. *J. Virol.* **1983**, *45* (2), 555–564.

(12) Mendelson, E.; Trempe, J. P.; Carter, B. J. Identification of the Trans-Acting Rep Proteins of Adeno-Associated Virus by Antibodies to a Synthetic Oligopeptide. *J. Virol.* **1986**, *60* (3), 823–832.

(13) Trempe, J. P.; Carter, B. J. Alternate MRNA Splicing Is Required for Synthesis of Adeno-Associated Virus VP1 Capsid Protein. *J. Virol.* **1988**, *62* (9), 3356–3363.

(14) Stutika, C.; Gogol-Döring, A.; Botschen, L.; Mietzsch, M.; Weger, S.; Feldkamp, M.; Chen, W.; Heilbronn, R. A Comprehensive RNA Sequencing Analysis of the Adeno-Associated Virus (AAV) Type 2 Transcriptome Reveals Novel AAV Transcripts, Splice Variants, and

Derived Proteins. *J. Virol.* **2016**, *90* (3), 1278–1289.

(15) Sonntag, F.; Schmidt, K.; Kleinschmidt, J. A. A Viral Assembly Factor Promotes AAV2 Capsid Formation in the Nucleolus. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107* (22), 10220–10225.

(16) Sonntag, F.; Köther, K.; Schmidt, K.; Weghofer, M.; Raupp, C.; Nieto, K.; Kuck, A.; Gerlach, B.; Böttcher, B.; Müller, O. J.; et al. The Assembly-Activating Protein Promotes Capsid Assembly of Different Adeno-Associated Virus Serotypes ▿. *J. Virol.* **2011**, *85* (23), 12686–12697.

(17) Maurer, A. C.; Pacouret, S.; Cepeda Diaz, A. K.; Blake, J.; Andres-Mateos, E.; Vandenberghe, L. H. The Assembly-Activating Protein Promotes Stability and Interactions between AAV's Viral Proteins to Nucleate Capsid Assembly. *Cell Rep.* **2018**, *23* (6), 1817–1830.

(18) Cao, M.; You, H.; Hermonat, P. L. The X Gene of Adeno-Associated Virus 2 (AAV2) Is Involved in Viral DNA Replication. *PLoS One* **2014**, *9* (8), e104596.

(19) Xie, Q.; Bu, W.; Bhatia, S.; Hare, J.; Somasundaram, T.; Azzi, A.; Chapman, M. S. The Atomic Structure of Adeno-Associated Virus (AAV-2), a Vector for Human Gene Therapy. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99* (16), 10405–10410.

(20) Ng, R.; Govindasamy, L.; Gurda, B. L.; McKenna, R.; Kozyreva, O. G.; Samulski, R. J.; Parent, K. N.; Baker, T. S.; Agbandje-McKenna, M. Structural Characterization of the Dual Glycan Binding Adeno-Associated Virus Serotype 6. *J. Virol.* **2010**, *84* (24), 12945–12957.

(21) DiMattia, M. A.; Nam, H.-J.; Van Vliet, K.; Mitchell, M.; Bennett, A.; Gurda, B. L.; McKenna, R.; Olson, N. H.; Sinkovits, R. S.; Potter, M.; et al. Structural Insight into the Unique Properties of Adeno-Associated Virus Serotype 9. *J. Virol.* **2012**, *86* (12), 6947–6958.

(22) Drouin, L. M.; Agbandje-McKenna, M. Adeno-Associated Virus Structural Biology as a Tool in Vector Development. *Future Virol.* **2013**, *8* (12), 1183–1199.

(23) Wu, Z.; Miller, E.; Agbandje-McKenna, M.; Samulski, R. J. Alpha2,3 and Alpha2,6 N-Linked Sialic Acids Facilitate Efficient Binding and Transduction by Adeno-Associated Virus Types 1 and 6. *J. Virol.* **2006**, *80* (18), 9093–9103.

(24) Summerford, C.; Samulski, R. J. Membrane-Associated Heparan Sulfate Proteoglycan Is a Receptor for Adeno-Associated Virus Type 2 Virions. *J. Virol.* **1998**, *72* (2), 1438–1445.

(25) Qing, K.; Mah, C.; Hansen, J.; Zhou, S.; Dwarki, V.; Srivastava, A. Human Fibroblast Growth Factor Receptor 1 Is a Co-Receptor for Infection by Adeno-Associated Virus 2.

Nat. Med. **1999**, *5* (1), 71–77.

(26) Kashiwakura, Y.; Tamayose, K.; Iwabuchi, K.; Hirai, Y.; Shimada, T.; Matsumoto, K.; Nakamura, T.; Watanabe, M.; Oshimi, K.; Daida, H. Hepatocyte Growth Factor Receptor Is a Coreceptor for Adeno-Associated Virus Type 2 Infection. *J. Virol.* **2005**, *79* (1), 609–614.

(27) Summerford, C.; Bartlett, J. S.; Samulski, R. J. AVβ5 Integrin: A Co-Receptor for Adeno-Associated Virus Type 2 Infection. *Nat. Med.* **1999**, *5* (1), 78–82.

(28) Asokan, A.; Hamra, J. B.; Govindasamy, L.; Agbandje-McKenna, M.; Samulski, R. J. Adeno-Associated Virus Type 2 Contains an Integrin A5β1 Binding Domain Essential for Viral Cell Entry. *J. Virol.* **2006**, *80* (18), 8961–8969.

(29) Akache, B.; Grimm, D.; Pandey, K.; Yant, S. R.; Xu, H.; Kay, M. A. The 37/67-Kilodalton Laminin Receptor Is a Receptor for Adeno-Associated Virus Serotypes 8, 2, 3, and 9. *J. Virol.* **2006**, *80* (19), 9831–9836.

(30) Handa, A.; Muramatsu, S. I.; Qiu, J.; Mizukami, H.; Brown, K. E. Adeno-Associated Virus (AAV)-3-Based Vectors Transduce Haematopoietic Cells Not Susceptible to Transduction with AAV-2-Based Vectors. *J. Gen. Virol.* **2000**, *81* (8), 2077–2084.

(31) Blackburn, S. D.; Steadman, R. A.; Johnson, F. B. Attachment of Adeno-Associated Virus Type 3H to Fibroblast Growth Factor Receptor 1. *Arch. Virol.* **2006**, *151* (3), 617–623.

(32) Ling, C.; Lu, Y.; Kalsi, J. K.; Jayandharan, G. R.; Li, B.; Ma, W.; Cheng, B.; Gee, S. W. Y.; McGoogan, K. E.; Govindasamy, L.; et al. Human Hepatocyte Growth Factor Receptor Is a Cellular Coreceptor for Adeno-Associated Virus Serotype 3. *Hum. Gene Ther.* **2010**, *21* (12), 1741–1747.

(33) Kaludov, N.; Brown, K. E.; Walters, R. W.; Zabner, J.; Chiorini, J. A. Adeno-Associated Virus Serotype 4 (AAV4) and AAV5 Both Require Sialic Acid Binding for Hemagglutination and Efficient Transduction but Differ in Sialic Acid Linkage Specificity. *J. Virol.* **2001**, *75* (15), 6884–6893.

(34) Walters, R. W.; Yi, S. M. P.; Keshavjee, S.; Brown, K. E.; Welsh, M. J.; Chiorini, J. A.; Zabner, J. Binding of Adeno-Associated Virus Type 5 to 2,3-Linked Sialic Acid Is Required for Gene Transfer. *J. Biol. Chem.* **2001**, *276* (23), 20610–20616.

(35) Di Pasquale, G.; Davidson, B. L.; Stein, C. S.; Martins, I.; Scudiero, D.; Monks, A.; Chiorini, J. A. Identification of PDGFR as a Receptor for AAV-5 Transduction. *Nat. Med.* **2003**, *9* (10), 1306–1312.

(36) Samulski, R. J.; Govindasamy, L.; Grieger, J. C.; Asokan, A.; Wu, Z.; Agbandje-McKenna,

M. Single Amino Acid Changes Can Influence Titer, Heparin Binding, and Tissue Tropism in Different Adeno-Associated Virus Serotypes. *J. Virol.* **2006**, *80* (22), 11393–11397.

(37) Bell, C. L.; Gurda, B. L.; Van Vliet, K.; Agbandje-McKenna, M.; Wilson, J. M. Identification of the Galactose Binding Domain of the Adeno-Associated Virus Serotype 9 Capsid. *J. Virol.* **2012**, *86* (13), 7326–7333.

(38) Kern, A.; Schmidt, K.; Leder, C.; Müller, O. J.; Wobus, C. E.; Bettinger, K.; Von der Lieth, C. W.; King, J. A.; Kleinschmidt, J. A. Identification of a Heparin-Binding Motif on Adeno-Associated Virus Type 2 Capsids. *J. Virol.* **2003**, *77* (20), 11072–11081.

(39) Opie, S. R.; Kenneth H. Warrington, J.; Agbandje-McKenna, M.; Zolotukhin, S.; Muzyczka, N. Identification of Amino Acid Residues in the Capsid Proteins of Adeno-Associated Virus Type 2 That Contribute to Heparan Sulfate Proteoglycan Binding. *J. Virol.* **2003**, *77* (12), 6995–7006.

(40) Pillay, S.; Meyer, N. L.; Puschnik, A. S.; Davulcu, O.; Diep, J.; Ishikawa, Y.; Jae, L. T.; Wosen, J. E.; Nagamine, C. M.; Chapman, M. S.; et al. An Essential Receptor for Adeno-Associated Virus Infection. *Nature* **2016**, *530* (7588), 108–112.

(41) Dudek, A. M.; Pillay, S.; Puschnik, A. S.; Nagamine, C. M.; Cheng, F.; Qiu, J.; Carette, J. E.; Vandenberghe, L. H. An Alternate Route for Adeno-Associated Virus (AAV) Entry Independent of AAV Receptor. *J. Virol.* **2018**, *92* (7), e02213-17.

(42) Bartlett, J. S.; Wilcher, R.; Samulski, R. J. Infectious Entry Pathway of Adeno-Associated Virus and Adeno-Associated Virus Vectors. *J. Virol.* **2000**, *74* (6), 2777–2785.

(43) Samulski, R.; Zhu, X.; Xiao, X.; Brook, J.; Housman, D.; Epstein, N.; Hunter, L. Targeted Integration of Adeno-Associated Virus (AAV) into Human Chromosome 19. *EMBO J.* **1991**, *10* (12), 3941.

(44) Kotin, R. M.; Linden, R. M.; Berns, K. I. Characterization of a Preferred Site on Human Chromosome 19q for Integration of Adeno-Associated Virus DNA by Non-Homologous Recombination. *EMBO J.* **1992**, *11* (13), 5071–5078.

(45) Linden, R. M.; Ward, P.; Giraud, C.; Winocour, E.; Berns, K. I. Site-Specific Integration by Adeno-Associated Virus. *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93* (21), 11288–11294.

(46) Wistuba, A.; Kern, A.; Weger, S.; Grimm, D.; Kleinschmidt, J. A. Subcellular Compartmentalization of Adeno-Associated Virus Type 2 Assembly. *J. Virol.* **1997**, *71* (2), 1341–1352.

(47) Matsushita, T.; Elliger, S.; Elliger, C.; Podsakoff, G.; Villarreal, L.; Kurtzman, G. J.; Iwaki, Y.; Colosi, P. Adeno-Associated Virus Vectors Can Be Efficiently Produced without Helper

Virus. *Gene Ther.* **1998**, *5* (7), 938–945.

(48) Xiao, X.; Li, J.; Samulski, R. J. Production of High-Titer Recombinant Adeno-Associated Virus Vectors in the Absence of Helper Adenovirus. *J. Virol.* **1998**, *72* (3), 2224–2232.

(49) Boutin, S.; Monteilhet, V.; Veron, P.; Leborgne, C.; Benveniste, O.; Montus, M. F.; Masurier, C. Prevalence of Serum IgG and Neutralizing Factors against Adeno-Associated Virus (AAV) Types 1, 2, 5, 6, 8, and 9 in the Healthy Population: Implications for Gene Therapy Using AAV Vectors. *Hum. Gene Ther.* **2010**, *21* (6), 704–712.

(50) Müller, O. J.; Kaul, F.; Weitzman, M. D.; Pasqualini, R.; Arap, W.; Kleinschmidt, J. A.; Trepel, M. Random Peptide Libraries Displayed on Adeno-Associated Virus to Select for Targeted Gene Therapy Vectors. *Nat. Biotechnol.* **2003**, *21* (9), 1040–1046.

(51) Michelfelder, S.; Trepel, M. *Adeno-Associated Viral Vectors and Their Redirection to Cell-Type Specific Receptors*, 1st ed.; Elsevier Inc., 2009; Vol. 67.

(52) Girod; Ried; Wobus; Lahm; Leike; Kleinschmidt; Deléage; Hallek. Genetic Capsid Modifications Allow Efficient Re-Targeting of Adeno-Associated Virus Type 2. *Nat. Med.* **1999**, *5* (12), 1438.

(53) Shi, W.; Bartlett, J. S. RGD Inclusion in VP3 Provides Adeno-Associated Virus Type 2 (AAV2)-Based Vectors with a Heparan Sulfate-Independent Cell Entry Mechanism. *Mol. Ther.* **2003**, *7* (4), 515–525.

(54) Büning, H.; Ried, M. U.; Perabo, L.; Gerner, F. M.; Huttner, N. A.; Enssle, J.; Hallek, M. Receptor Targeting of Adeno-Associated Virus Vectors. *Gene Ther.* **2003**, *10* (14), 1142–1151.

(55) Opie, S. R.; Warrington, K. H.; Agbandje-McKenna, M.; Zolotukhin, S.; Muzyczka, N. Identification of Amino Acid Residues in the Capsid Proteins of Adeno-Associated Virus Type 2 That Contribute to Heparan Sulfate Proteoglycan Binding. *J. Virol.* **2003**, *77* (12), 6995–7006.

(56) Wu, Z.; Asokan, A.; Grieger, J. C.; Govindasamy, L.; Agbandje-McKenna, M.; Samulski, R. J. Single Amino Acid Changes Can Influence Titer, Heparin Binding, and Tissue Tropism in Different Adeno-Associated Virus Serotypes. *J. Virol.* **2006**, *80* (22), 11393–11397.

(57) Ried, M. U.; Girod, A.; Leike, K.; Büning, H.; Hallek, M. Adeno-Associated Virus Capsids Displaying Immunoglobulin-Binding Domains Permit Antibody-Mediated Vector Retargeting to Specific Cell Surface Receptors. *J. Virol.* **2002**, *76* (9), 4559–4566.

(58) Liu, Y.; Fang, Y.; Zhou, Y.; Zandi, E.; Lee, C.-L.; Joo, K.-I.; Wang, P. Site-Specific

Modification of Adeno-Associated Viruses via a Genetically Engineered Aldehyde Tag. *Small* **2013**, *9* (3), 421–429.

(59) Münch, R. C.; Janicki, H.; Völker, I.; Rasbach, A.; Hallek, M.; Büning, H.; Buchholz, C. J. Displaying High-Affinity Ligands on Adeno-Associated Viral Vectors Enables Tumor Cell-Specific and Safe Gene Transfer. *Mol. Ther.* **2013**, *21* (1), 109–118.

(60) Hagen, S.; Baumann, T.; Wagner, H. J.; Morath, V.; Kaufmann, B.; Fischer, A.; Bergmann, S.; Schindler, P.; Arndt, K. M.; Müller, K. M. Modular Adeno-Associated Virus (RAAV) Vectors Used for Cellular Virus-Directed Enzyme Prodrug Therapy. *Sci. Rep.* **2014**, *4*.

(61) Michelfelder, S.; Varadi, K.; Raupp, C.; Hunger, A.; Körbelin, J.; Pahrmann, C.; Schrepfer, S.; Müller, O. J.; Kleinschmidt, J. A.; Trepel, M.; et al. Peptide Ligands Incorporated into the Threefold Spike Capsid Domain to Re-Direct Gene Transduction of AAV8 and AAV9 in Vivo. *PLoS One* **2011**, *6* (8).

(62) Sayroo, R.; Nolasco, D.; Yin, Z.; Colon-Cortes, Y.; Pandya, M.; Ling, C.; Aslanidi, G. Development of Novel AAV Serotype 6 Based Vectors with Selective Tropism for Human Cancer Cells. *Gene Ther.* **2016**, *23* (1), 18–25.

(63) Rajendran, S.; Collins, S.; Van Pijkeren, J. P.; O'Hanlon, D.; O'Sullivan, G. C.; Tangney, M. Targeting of Breast Metastases Using a Viral Gene Vector with Tumour-Selective Transcription. *Anticancer Res.* **2011**, *31* (5), 1627–1635.

(64) Yang, X.; Rutnam, Z. J.; Jiao, C.; Wei, D.; Xie, Y.; Du, J.; Zhong, L.; Yang, B. B. An Anti-Let-7 Sponge Decoys and Decays Endogenous Let-7 Functions. *Cell Cycle* **2012**, *11* (16), 3097–3108.

(65) Pasquinelli, A. E. The Primary Target of Let-7 MicroRNA. *Biochem. Soc. Trans.* **2013**, *41* (4), 821–824.

(66) Xie, J.; Burt, D. R.; Gao, G. Adeno-Associated Virus-Mediated MicroRNA Delivery and Therapeutics. *Semin. Liver Dis.* **2015**, *35* (1), 81–88.

(67) Huber, B. E.; Richards, C. A.; Austin, E. A. Virus-Directed Enzyme/Prodrug Therapy (VDEPT). Selectively Engineering Drug Sensitivity into Tumors. *Ann. N. Y. Acad. Sci.* **1994**, *716* (14), 104–114; discussion 140-3.

(68) MF, P.; HJ, L. EGFR, HER2 and VEGF Pathways: Validated Targets for Cancer Treatment. *Drugs* **2007**, *67* (14), 2045–2075.

(69) Roskoski, R. The ErbB/HER Family of Protein-Tyrosine Kinases and Cancer. *Pharmacol. Res.* **2014**, *79*, 34–74.

(70) Yarden, Y.; Sliwkowski, M. X. Untangling the ErbB Signalling Network. *Nat. Rev. Mol.*

Cell Biol. **2001**, *2* (2), 127–137.

(71) Wilson, K. J.; Gilmore, J. L.; Foley, J.; Lemmon, M. A.; Riese, D. J. Functional Selectivity of EGF Family Peptide Growth Factors: Implications for Cancer. *Pharmacol. Ther.* **2009**, *122* (1), 1–8.

(72) Friedländer, E.; Barok, M.; Szöllosi, J.; Vereb, G. ErbB-Directed Immunotherapy: Antibodies in Current Practice and Promising New Agents. *Immunol. Lett.* **2008**, *116* (2), 126–140.

(73) Nicholson, R. I.; Gee, J. M. W.; Harper, M. E. EGFR and Cancer Prognosis. *Eur. J. Cancer* **2001**, *37*, S9–S15.

(74) Feiner, R. C.; Müller, K. M. Recent Progress in Protein-Protein Interaction Study for EGFR-Targeted Therapeutics. *Expert Rev. Proteomics* **2016**, *13* (9), 817–832.

(75) Masui, H.; Kawamoto, T.; Sato, J. D.; Wolf, B.; Sato, G.; Mendelsohn, J. Growth Inhibition of Human Tumor Cells in Athymic Mice by Anti- Epidermal Growth Factor Receptor Monoclonal Antibodies. *Cancer Res* **1984**, *44* (3), 1002–1007.

(76) Goldstein, N. I.; Prewett, M.; Zuklys, K.; Rockwell, P.; Mendelsohn, J. Biological Efficacy of a Chimeric Antibody to the Epidermal Growth Factor Receptor in a Human Tumor Xenograft Model. *Clin. Cancer Res.* **1995**, *1* (11), 1311–1318.

(77) Kenanova, V.; Olafsen, T.; Crow, D. M.; Sundaresan, G.; Subbarayan, M.; Carter, N. H.; Ikle, D. N.; Yazaki, P. J.; Chatziioannou, A. F.; Gambhir, S. S.; et al. Tailoring the Pharmacokinetics and Positron Emission Tomography Imaging Properties of Anti-Carcinoembryonic Antigen Single-Chain Fv-Fc Antibody Fragments. *Cancer Res.* **2005**, *65* (2), 622–631.

(78) Shah, D. K. Pharmacokinetic and Pharmacodynamic Considerations for the next Generation Protein Therapeutics. *J. Pharmacokinet. Pharmacodyn.* **2015**, *42* (5), 553–571.

(79) Forrer, P.; Stumpp, M. T.; Binz, H. K.; Plückthun, A. A Novel Strategy to Design Binding Molecules Harnessing the Modular Nature of Repeat Proteins. *FEBS Lett.* **2003**, *539* (1–3), 2–6.

(80) Boersma, Y. L.; Chao, G.; Steiner, D.; Wittrup, K. D.; Plückthun, A. Bispecific Designed Ankyrin Repeat Proteins (DARPins) Targeting Epidermal Growth Factor Receptor Inhibit A431 Cell Proliferation and Receptor Recycling. *J. Biol. Chem.* **2011**, *286* (48), 41273–41285.

(81) Li, Z.; Zhao, R.; Wu, X.; Sun, Y.; Yao, M.; Li, J.; Xu, Y.; Gu, J. Identification and Characterization of a Novel Peptide Ligand of Epidermal Growth Factor Receptor for

Targeted Delivery of Therapeutics. *FASEB J.* **2005**, *19* (14), 1978–1985.

(82) Li, Z.; Zhao, R.; Wu, X.; Sun, Y.; Yao, M.; Li, J.; Xu, Y.; Gu, J. Identification and Characterization of a Novel Peptide Ligand of Epidermal Growth Factor Receptor for Targeted Delivery of Therapeutics. *FASEB J.* **2005**, *19* (14), 1978–1985.

(83) Kuo, W.-T.; Lin, W.-C.; Chang, K.-C.; Huang, J.-Y.; Yen, K.-C.; Young, I.-C.; Sun, Y.-J.; Lin, F.-H. Quantitative Analysis of Ligand-EGFR Interactions: A Platform for Screening Targeting Molecules. *PLoS One* **2015**, *10* (2), e0116610.

(84) Williams, T. M.; Sable, R.; Singh, S.; Vicente, M. G. H.; Jois, S. D. Peptide Ligands for Targeting the Extracellular Domain of EGFR: Comparison between Linear and Cyclic Peptides. *Chem. Biol. Drug Des.* **2018**, *91* (2), 605–619.

(85) Hanold, L. E.; Oruganty, K.; Ton, N. T.; Beedle, A. M.; Kannan, N.; Kennedy, E. J. Inhibiting EGFR Dimerization Using Triazolyl-Bridged Dimerization Arm Mimics. *PLoS One* **2015**, *10* (3), e0118796.

(86) Toyama, K.; Mizuguchi, T.; Nomura, W.; Tamamura, H. Functional Evaluation of Fluorescein-Labeled Derivatives of a Peptide Inhibitor of the EGF Receptor Dimerization. *Bioorg. Med. Chem.* **2016**, *24* (16), 3406–3412.

(87) Abourbeh, G.; Shir, A.; Mishani, E.; Ogris, M.; Rödl, W.; Wagner, E.; Levitzki, A. PolyIC GE11 Polyplex Inhibits EGFR-Overexpressing Tumors. *IUBMB Life* **2012**, *64* (4), 324–330.

(88) Toyama, K.; Kobayakawa, T.; Nomura, W.; Tamamura, H. Inhibition of EGFR Activation by Bivalent Ligands Based on a Cyclic Peptide Mimicking the Dimerization Arm Structure of EGFR. *Chem. Pharm. Bull.* **2018**, *66* (11), 1083–1089.

(89) Master, A. M.; Sen Gupta, A. EGF Receptor-Targeted Nanocarriers for Enhanced Cancer Treatment. *Nanomedicine* **2012**, *7* (12), 1895–1906.

(90) Xu, W. W.; Liu, D. Y.; Cao, Y. C.; Wang, X. Y. GE11 Peptide-Conjugated Nanoliposomes to Enhance the Combinational Therapeutic Efficacy of Docetaxel and SiRNA in Laryngeal Cancers. *Int. J. Nanomedicine* **2017**, *12*, 6461–6470.

(91) Cheng, L.; Huang, F. Z.; Cheng, L. F.; Zhu, Y. Q.; Hu, Q.; Li, L.; Wei, L.; Chen, D. W. GEII-Modified Liposomes for Non-Small Cell Lung Cancer Targeting: Preparation, Ex Vitro and in Vivo Evaluation. *Int. J. Nanomedicine* **2014**, *9* (1), 921–935.

(92) Chariou, P. L.; Lee, K. L.; Wen, A. M.; Gulati, N. M.; Stewart, P. L.; Steinmetz, N. F. Detection and Imaging of Aggressive Cancer Cells Using an Epidermal Growth Factor Receptor (EGFR)-Targeted Filamentous Plant Virus-Based Nanoparticle. *Bioconjug.*

Chem. **2015**, *26* (2), 262–269.

(93) Körbelin, J.; Sieber, T.; Michelfelder, S.; Lunding, L.; Spies, E.; Hunger, A.; Alawi, M.; Rapti, K.; Indenbirken, D.; Müller, O. J.; et al. Pulmonary Targeting of Adeno-Associated Viral Vectors by Next-Generation Sequencing-Guided Screening of Random Capsid Displayed Peptide Libraries. *Mol. Ther.* **2016**, *24* (6), 1050–1061.

(94) Yu, C.-Y.; Yuan, Z.; Cao, Z.; Wang, B.; Qiao, C.; Li, J.; Xiao, X. A Muscle-Targeting Peptide Displayed on AAV2 Improves Muscle Tropism on Systemic Delivery. *Gene Ther.* **2009**, *16* (8), 953–962.

(95) Jiang, H.; Zhang, J.; Shi, B. Z.; Xu, Y. H.; Li, Z. H.; Gu, J. R. Application of EGFP-EGF Fusions to Explore Mechanism of Endocytosis of Epidermal Growth Factor. *Acta Pharmacol. Sin.* **2007**, *28* (1), 111–117.

(96) Gross, L. A.; Baird, G. S.; Hoffman, R. C.; Baldridge, K. K.; Tsien, R. Y. The Structure of the Chromophore within DsRed, a Red Fluorescent Protein from Coral. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97* (22), 11990–11995.

(97) Berkmen, M. Production of Disulfide-Bonded Proteins in Escherichia Coli. *Protein Expr. Purif.* **2012**, *82* (1), 240–251.

(98) Borrell, K. L.; Cancglin, C.; Stinger, B. L.; Defrates, K. G.; Caputo, G. A.; Wu, C.; Vaden, T. D. An Experimental and Molecular Dynamics Study of Red Fluorescent Protein MCherry in Novel Aqueous Amino Acid Ionic Liquids. *J. Phys. Chem. B* **2017**, *121* (18), 4823–4832.

(99) Baier, T.; Kros, D.; Feiner, R. C.; Lauersen, K. J.; Müller, K. M.; Kruse, O. Engineered Fusion Proteins for Efficient Protein Secretion and Purification of a Human Growth Factor from the Green Microalga Chlamydomonas Reinhardtii. *ACS Synth. Biol.* **2018**, *7* (11), 2547–2557.

(100) Dickson, R. B.; Hanover, J. A.; Willingham, M. C.; Pastan, I. Prelysosomal Divergence of Transferrin and Epidermal Growth Factor during Receptor-Mediated Endocytosis. *Biochemistry* **1983**, *22* (24), 5667–5674.

(101) Reilly, R. M.; Kiarash, R.; Sandhu, J.; Lee, Y. W.; Cameron, R. G.; Hendler, A.; Vallis, K.; Gariépy, J. A Comparison of EGF and MAb 528 Labeled with 111In for Imaging Human Breast Cancer. *J. Nucl. Med.* **2000**, *41* (5), 903–911.

(102) Turner, R. A.; Oliver, A. G.; Lokey, R. S. Click Chemistry as a Macrocyclization Tool in the Solid-Phase Synthesis of Small Cyclic Peptides. *Org. Lett.* **2007**, *9* (24), 5011–5014.

(103) Cantel, S.; Isaad, A. L. C.; Scrima, M.; Levy, J. J.; DiMarchi, R. D.; Rovero, P.; Halperin, J. A.; D'Ursi, A. M.; Papini, A. M.; Chorev, M. Synthesis and Conformational Analysis of

a Cyclic Peptide Obtained via i to I+4 Intramolecular Side-Chain to Side-Chain Azide-Alkyne 1,3-Dipolar Cycloaddition. *J. Org. Chem.* **2008**, *73* (15), 5663–5674.

(104) Haigler, H.; Ash, J. F.; Singer, S. J.; Cohen, S. Visualization by Fluorescence of the Binding and Internalization of Epidermal Growth Factor in Human Carcinoma Cells A-431. *Proc. Natl. Acad. Sci. U. S. A.* **1978**, *75* (7), 3317–3321.

(105) Vieira, A. V; Lamaze, C.; Schmid, S. L. Control of EGF Receptor Signaling by Clathrin-Mediated Endocytosis. *Science* **1996**, *274* (5295), 2086–2089.

(106) Rossi, A. M.; Taylor, C. W. Analysis of Protein-Ligand Interactions by Fluorescence Polarization. *Nat. Protoc.* **2011**, *6* (3), 365–387.

(107) Moerke, N. J. Fluorescence Polarization (FP) Assays for Monitoring Peptide-Protein or Nucleic Acid-Protein Binding. *Curr. Protoc. Chem. Biol.* **2009**, *1* (December), 1–15.

(108) Roehrl, M. H. A.; Wang, J. Y.; Wagner, G. A General Framework for Development and Data Analysis of Competitive High-Throughput Screens for Small-Molecule Inhibitors of Protein-Protein Interactions by Fluorescence Polarization. *Biochemistry* **2004**, *43* (51), 16056–16066.

(109) Lemmon, M. A.; Bu, Z.; Ladbury, J. E.; Zhou, M.; Pinchasi, D.; Lax, I.; Engelman, D. M.; Schlessinger, J. Two EGF Molecules Contribute Additively to Stabilization of the EGFR Dimer. *EMBO J.* **1997**, *16* (2), 281–294.

(110) Bessman, N. J.; Bagchi, A.; Ferguson, K. M.; Lemmon, M. A. Complex Relationship between Ligand Binding and Dimerization in the Epidermal Growth Factor Receptor. *Cell Rep.* **2014**, *9* (4), 1306–1317.

(111) Wu, P.; Xiao, W.; Conlon, T.; Hughes, J.; Agbandje-McKenna, M.; Ferkol, T.; Flotte, T.; Muzyczka, N. Mutational Analysis of the Adeno-Associated Virus Type 2 (AAV2) Capsid Gene and Construction of AAV2 Vectors with Altered Tropism. *J. Virol.* **2000**, *74* (18), 8635–8647.

(112) Ellis, B. L.; Hirsch, M. L.; Barker, J. C.; Connelly, J. P.; Steininger, R. J.; Porteus, M. H. A Survey of Ex Vivo/in Vitro Transduction Efficiency of Mammalian Primary Cells and Cell Lines with Nine Natural Adeno-Associated Virus (AAV1-9) and One Engineered Adeno-Associated Virus Serotype. *Virol. J.* **2013**, *10* (1), 74.

(113) Ran, F. A.; Hsu, P. D.; Wright, J.; Agarwala, V.; Scott, D. A.; Zhang, F. Genome Engineering Using the CRISPR-Cas9 System. *Nat. Protoc.* **2013**, *8* (11), 2281–2308.

(114) Summerford, C.; Samulski, R. J. AAVR: A Multi-Serotype Receptor for AAV. *Mol. Ther.* **2016**, *24* (4), 663–666.

(115) Lokman, N. A.; Elder, A. S. F.; Ricciardelli, C.; Oehler, M. K. Chick Chorioallantoic Membrane (CAM) Assay as an in Vivo Model to Study the Effect of Newly Identified Molecules on Ovarian Cancer Invasion and Metastasis. *Int. J. Mol. Sci.* **2012**, *13* (8), 9959–9970.

(116) Durand, J.-O.; Jimenez, C. M.; Tamanoi, F.; Khashab, N.; Croissant, J.; Ratliff, L.; Shahin, S. A.; Vu, B. T.; Yik, T.; Raehm, L.; et al. Chick Chorioallantoic Membrane Assay as an in Vivo Model to Study the Effect of Nanoparticle-Based Anticancer Drugs in Ovarian Cancer. *Sci. Rep.* **2018**, *8* (1), 1–10.

(117) Pathak, R. R.; Li, M.; Aguirre-Ghiso, J. A.; Sikora, A. G.; Lopez-Rivera, E.; Friedman, S. L. The In Ovo </Em>Chick Chorioallantoic Membrane (CAM) Assay as an Efficient Xenograft Model of Hepatocellular Carcinoma. *J. Vis. Exp.* **2015**, No. 104, 1–6.

(118) Valdes, T. I.; Kreutzer, D.; Moussy, F. The Chick Chorioallantoic Membrane as a Novel in Vivo Model for the Testing of Biomaterials. *J. Biomed. Mater. Res.* **2002**, *62* (2), 273–282.

(119) Zincarelli, C.; Soltys, S.; Rengo, G.; Rabinowitz, J. E. Analysis of AAV Serotypes 1-9 Mediated Gene Expression and Tropism in Mice after Systemic Injection. *Mol. Ther.* **2008**, *16* (6), 1073–1080.

(120) Duan, D. Systemic Delivery of Adeno-Associated Viral Vectors. *Curr. Opin. Virol.* **2016**, *21*, 16–25.

(121) Wu, Z.; Asokan, A.; Samulski, R. J. Adeno-Associated Virus Serotypes: Vector Toolkit for Human Gene Therapy. *Mol. Ther.* **2006**, *14* (3), 316–327.

(122) Merkel, S. F.; Andrews, A. M.; Lutton, E. M.; Mu, D.; Hudry, E.; Hyman, B. T.; Maguire, C. A.; Ramirez, S. H. Trafficking of AAV Vectors across a Model of the Blood-Brain Barrier; a Comparative Study of Transcytosis and Transduction Using Primary Human Brain Endothelial Cells. *J. Neurochem.* **2016**, 216–230.

(123) Weber-Adrian, D.; Heinen, S.; Silburt, J.; Noroozian, Z.; Aubert, I. The Human Brain Endothelial Barrier: Transcytosis of AAV9, Transduction by AAV2. *J. Neurochem.* **2016**, *140* (2), 192–194.

(124) Bigland, C. H.; Triantaphyllopoulos, D. C. A Re-Evaluation of the Clotting Time of Chicken Blood. *Nature* **1960**, *186* (1950), 644.

(125) Wechsler, W. Die Entwicklung Der Gefäße Und Pervivasculären Gewebsräume Im Zentralnervensystem von Hühnern. *Z. Anat. Entwicklungsgesch.* **1965**, *124* (4), 367–395.

(126) Bai, H.; Liu, Q.; Shi, M.; Zhang, J. Erlotinib and Gefitinib Treatments of the Lung Cancer in an Elderly Patient Result in Gastrointestinal Bleeding. *Pakistan J. Med. Sci.* **2013**, *29* (5),

1278–1279.

(127) Duan, S. J.; Gao, Z. M.; Wang, P. L.; Gong, B. C.; Huang, H. W.; Luo, L.; Wang, X.; Xing, Y. N.; Xu, H. M.; Liu, F. N. Rare Upper Gastrointestinal Hemorrhage of Cetuximab. *Med. (United States)* **2017**, *96* (51), 1–4.

(128) iGEM-Team-Freiburg. Virus Construction Kit for Therapy. *http://2010.igem.org/Team:Freiburg_Bioware* **2010**.

(129) Zolotukhin, S.; Byrne, B. J.; Mason, E.; Zolotukhin, I.; Potter, M.; Chesnut, K.; Summerford, C.; Samulski, R. J.; Muzyczka, N. Recombinant Adeno-Associated Virus Purification Using Novel Methods Improves Infectious Titer and Yield. *Gene Ther.* **1999**, *6* (6), 973–985.

(130) Guo, P.; El-Gohary, Y.; Prasadan, K.; Shiota, C.; Xiao, X.; Wiersch, J.; Paredes, J.; Tulachan, S.; Gittes, G. K. Rapid and Simplified Purification of Recombinant Adeno-Associated Virus. *J. Virol. Methods* **2012**, *183* (2), 139–146.

(131) Horowitz, E. D.; Rahman, K. S.; Bower, B. D.; Dismuke, D. J.; Falvo, M. R.; Griffith, J. D.; Harvey, S. C.; Asokan, A. Biophysical and Ultrastructural Characterization of Adeno-Associated Virus Capsid Uncoating and Genome Release. *J. Virol.* **2013**, *87* (6), 2994–3002.

(132) Zeltner, N.; Kohlbrenner, E.; Clément, N.; Weber, T.; Linden, R. M. Near-Perfect Infectivity of Wild-Type AAV as Benchmark for Infectivity of Recombinant AAV Vectors. *Gene Ther.* **2010**, *17* (7), 872–879.

(133) Bernaud, J.; Rossi, A.; Fis, A.; Gardette, L.; Aillot, L.; Büning, H.; Castelnovo, M.; Salvetti, A.; Faivre-Moskalenko, C. Characterization of AAV Vector Particle Stability at the Single-Capsid Level. *J. Biol. Phys.* **2018**, *44* (2), 181–194.

(134) Bennett, A.; Patel, S.; Mietzsch, M.; Jose, A.; Lins-Austin, B.; Yu, J. C.; Bothner, B.; McKenna, R.; Agbandje-McKenna, M. Thermal Stability as a Determinant of AAV Serotype Identity. *Mol. Ther. - Methods Clin. Dev.* **2017**, *6* (September), 171–182.

(135) Rayaprolu, V.; Kruse, S.; Kant, R.; Venkatakrishnan, B.; Movahed, N.; Brooke, D.; Lins, B.; Bennett, A.; Potter, T.; McKenna, R.; et al. Comparative Analysis of Adeno Associated Virus Capsid Stability and Dynamics. *J. Virol.* **2013**, *87* (24), 13150–13160.

(136) Bennett, A. D.; Wong, K.; Lewis, J.; Tseng, Y. S.; Smith, J. K.; Chipman, P.; McKenna, R.; Samulski, R. J.; Kleinschmidt, J.; Agbandje-McKenna, M. AAV6 K531 Serves a Dual Function in Selective Receptor and Antibody ADK6 Recognition. *Virology* **2018**, *518* (December 2017), 369–376.

(137) Kotin, R. Large-Scale Recombinant Adeno-Associated Virus Production. *Hum. Mol. Genet.* **2011**, 1–23.

(138) Lin, Y.-C.; Boone, M.; Meuris, L.; Lemmens, I.; Van Roy, N.; Soete, A.; Reumers, J.; Moisse, M.; Plaisance, S.; Drmanac, R.; et al. Genome Dynamics of the Human Embryonic Kidney 293 Lineage in Response to Cell Biology Manipulations. *Nat. Commun.* **2014**, *5* (11), 4767.

(139) Louis, N.; Evelegh, C.; Graham, F. L. Cloning and Sequencing of the Cellular-Viral Junctions from the Human Adenovirus Type 5 Transformed 293 Cell Line. *Virology* **1997**, *233* (2), 423–429.

(140) Graham, F. L.; Smiley, J.; Russell, W. C.; Nairn, R. Characteristics of a Human Cell Line Transformed by DNA from Human Adenovirus Type 5. *J. Gen. Virol.* **1977**, *36* (1), 59–74.

(141) Xiao, X.; Li, J.; Samulski, R. J. Production of High-Titer Recombinant Adeno-Associated Virus Vectors in the Absence of Helper Adenovirus. *J. Virol.* **1998**, *72* (3), 2224–2232.

(142) Ganesan, L. P.; Mohanty, S.; Kim, J.; Clark, K. R.; Robinson, J. M.; Anderson, C. L. Rapid and Efficient Clearance of Blood-Borne Virus by Liver Sinusoidal Endothelium. *PLoS Pathog.* **2011**, *7* (9), e1002281.

(143) Mohammadi, E. S.; Ketner, E. A.; Johns, D. C.; Ketner, G. Expression of the Adenovirus E4 34k Oncoprotein Inhibits Repair of Double Strand Breaks in the Cellular Genome of a 293-Based Inducible Cell Line. *Nucleic Acids Res.* **2004**, *32* (8), 2652–2659.

(144) Berk, A. J. Adenovirus Promoters and E1A Transactivation. *Annu. Rev. Genet.* **1986**, *20* (1), 45–79.

(145) Klessig, D. F.; Grodzicker, T.; Cleghon, V. Construction of Human Cell Lines Which Contain and Express the Adenovirus DNA Binding Protein Gene by Cotransformation with the HSV-1 Tk Gene. *Virus Res.* **1984**, *1* (2), 169–188.

(146) Grifman, M.; Chen, N. N.; Gao, G. p; Cathomen, T.; Wilson, J. M.; Weitzman, M. D. Overexpression of Cyclin A Inhibits Augmentation of Recombinant Adeno-Associated Virus Transduction by the Adenovirus E4orf6 Protein. *J. Virol.* **1999**, *73* (12), 10010–10019.

(147) Koerber, J. T.; Zabner, J.; Dickey, D. D.; Keshavjee, S.; Schaffer, D. V.; Kaspar, B. K.; Excoffon, K. J. D. A.; Murtha, M. Directed Evolution of Adeno-Associated Virus to an Infectious Respiratory Virus. *Proc. Natl. Acad. Sci.* **2009**, *106* (10), 3865–3870.

(148) Asuri, P.; Bartel, M. A.; Vazin, T.; Jang, J. H.; Wong, T. B.; Schaffer, D. V. Directed Evolution of Adeno-Associated Virus for Enhanced Gene Delivery and Gene Targeting in

Human Pluripotent Stem Cells. *Mol. Ther.* **2012**, *20* (2), 329–338.

(149) Stahl, P. J.; Cruz, J. C.; Li, Y.; Michael Yu, S.; Hristova, K. On-the-Resin N-Terminal Modification of Long Synthetic Peptides. *Anal. Biochem.* **2012**, *424* (2), 137–139.

(150) Pohl, M.; Ambrosius, D.; Grötzinger, J.; Kretzschmar, T.; Saunders, D.; Wollmer, A.; Brandenburg, D.; Bitter-Suermann, D.; Höcker, H. Cyclic Disulfide Analogues of the Complement Component C3a. Synthesis and Conformational Investigations. *Int. J. Pept. Protein Res.* **1993**, *41* (4), 362–375.

(151) Guo, P.; El-Gohary, Y.; Prasadan, K.; Shiota, C.; Xiao, X.; Wiersch, J.; Paredes, J.; Tulachan, S.; Gittes, G. K. Rapid and Simplified Purification of Recombinant Adeno-Associated Virus. *J. Virol. Methods* **2012**, *183* (2), 139–146.

(152) Schindelin, J.; Rueden, C. T.; Hiner, M. C.; Eliceiri, K. W. The ImageJ Ecosystem: An Open Platform for Biomedical Image Analysis. *Mol. Reprod. Dev.* **2017**, *82* (7–8), 518–529.

(153) Crosson, S. M.; Dib, P.; Smith, J. K.; Zolotukhin, S. Helper-Free Production of Laboratory Grade AAV and Purification by Iodixanol Density Gradient Centrifugation. *Mol. Ther. Methods Clin. Dev.* **2018**, *10* (September), 1–7.

9. Appendix: Publications

9 Appendix: Publications

Title:

Recent progress in protein-protein interaction study for

EGFR-targeted therapeutics

Authors:

Rebecca C. Feiner, Kristian M- Müller

This is the authors accepted manuscript of an article published as the version of record in Expert Review of Proteomics © 2016 republished by permission of Informa UK Limited, trading as Taylor & Francis Group, link to this article: https://doi.org/10.1080/14789450.2016.1212665.

9 Appendix: Publications

Expert Review of Proteomics

ISSN: 1478-9450 (Print) 1744-8387 (Online) Journal homepage: http://www.tandfonline.com/loi/ieru20

Recent progress in protein-protein interaction study for EGFR-targeted therapeutics

Rebecca Christine Feiner & Kristian Mark Müller

To cite this article: Rebecca Christine Feiner & Kristian Mark Müller (2016): Recent progress in protein-protein interaction study for EGFR-targeted therapeutics, Expert Review of Proteomics, DOI: 10.1080/14789450.2016.1212665

To link to this article: http://dx.doi.org/10.1080/14789450.2016.1212665

Accepted author version posted online: 18 Jul 2016.
Published online: 18 Jul 2016.

Submit your article to this journal

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
http://www.tandfonline.com/action/journalInformation?journalCode=ieru20

Publisher: Taylor & Francis

Journal: *Expert Review of Proteomics*

DOI: 10.1080/14789450.2016.1212665

Review

Recent progress in protein-protein interaction study for EGFR-targeted therapeutics

Feiner, Rebecca Christine

Universitat Bielefeld Ringgold standard institution - Technische Fakultät

Universitätsstr. 25 Bielefeld, Bielefeld, Nordrhein-Westfalen 33501

Germany

Müller, Kristian Mark

Corresponding Author - Kristian@syntbio.net

Universitat Bielefeld Ringgold standard institution - Technische Fakultät

Universitätstr. 25, Bielefeld, Nordrhein-Westfalen 33501

Germany

ACCEPTED MANUSCRIPT

Abstract

Introduction:

Epidermal growth factor receptor (EGFR) expression is upregulated in many tumors and its aberrant signaling drives progression of many cancer types. Consequently, EGFR has become a clinically validated target as extracellular tumor marker for antibodies as well as for tyrosine kinase inhibitors. Within the last years, new mechanistic insights were uncovered and, based on clinical experience as well as progress in protein engineering, novel bio-therapeutic approaches were developed and tested.

Areas covered:

The potential therapeutic targeting arsenal in the fight against cancer now encompasses bispecific or biparatopic antibodies, DARPins, Adnectins, Affibodies, peptides and combinations of these binding molecules with viral- and nano-particles. We review past and recent binding proteins from the literature and include a brief description of the various targeting approaches. Special attention is given to the binding modes with the EGFR.

Expert Commentary:

Clinical data from the three approved anti EGFR antibodies indicate that there is room for improved therapeutic efficacy. Having choices in size, affinity, avidity and the mode of EGFR binding as well as the possibility to combine various effector functions opens the possibility to rationally design more effective therapeutics.

ACCEPTED MANUSCRIPT

Keywords

erbB1, EGF-R, tumor targeting, cancer targeting, antibody, antigen-binding scaffolds, virus-like particles

1 Introduction to the ErbB family of tyrosine kinases

1.1 ErbB tyrosine kinases

ErbB receptors are a well-studied receptor protein kinase (RTK) family comprising four members: the epidermal growth factor receptor (EGFR/ErbB1) as well as the closely related Her2 (ErbB2/NEU), Her3 (ErbB3) and Her4 (ErbB4). As type 1 transmembrane tyrosine kinases they play important roles in cell signaling affecting processes related to cell growth and development, differentiation, cell cycle progression, apoptosis and transcription [1,2]. They are expressed ubiquitously in epithelial, mesenchymal and neuronal cells and their progenitors. The gene name *erbB* originates from the name of the erythroblastic leukemia viral oncogene, which is homologous to the receptors [3]. Based on early analysis of the primary amino acid sequence with 1186 residues Ulrich et al. suggested that the receptor has an extracellular domain, a single hydrophobic α-helical transmembrane domain, and an intracellular domain with a protein kinase activity [4]. This picture was later detailed by crystal structure analyses [5]. The extracellular domain mediates ligand binding as well as homo- and heterodimerization of the ErbB family members. Dimerization upon ligand binding leads to conformational changes that are transmitted to the intracellular juxtamembrane helix and result in the activation of the kinase domain [6]. The kinase then phosphorylates one or more tyrosine residues in the neighboring RTKs which serve as docking sites for the assembly and activation of intracellular signaling proteins [1].

ACCEPTED MANUSCRIPT

1.2 Structure of ErbB extracellular domains

ErbB receptors consist of an N-terminal extracellular domain, which has in the case of EGFR a size of 622 amino acids. This domain can be divided in four subdomains I-IV. Domains I and III are leucine-rich segments and contact the ligands, whereas domains II and IV are cysteine rich and contain several disulfide bonds. Domain II is of special interest, because it mediates ErbB receptor homo- and heterodimerization via the exchange of a so called dimerization arm [2,7].

1.3 EGFR ectodomain autoinhibition

Once domain I and III of the EGFR sandwich one of the natural ligands, the receptor is stabilized in the extended conformation [8]. As a result the dimerization arm of domain II becomes exposed promoting dimerization and subsequent signalling [9]. For the inactive state crystallographic data suggest a resting state, the so called auto-tethered form, in which domain II is completely occluded by intramolecular interaction with domain IV (Figure 1). This auto-tethered conformation represents a monomeric form that is unable to dimerize [10]. In experiments an equilibrium of receptor monomers to ligand-free receptor dimers was found on living cells promoting the idea of preformed dimers prior to ligand binding, which may originate from non-tethered conformations [11]. Molecular dynamic studies indeed suggest that the extended monomer is not stable, a compact conformation is most likely, and that the receptor might be able to form ligand-free dimers with the need for autoinhibition in healthy cells [12]. Many cellular and biophysical studies of ErbB ligand binding and activation are limited to one specific cell type or model system making direct comparisons or generalizations difficult. In nature, a plethora of combinations of receptor ligands, homo- and heteromers and cellular factors such as receptor densities on the cell surface, cell membrane composition and availability of accessory proteins exists, mounting to a continuing challenge to pinpoint precise inactive and active states. Even activation with a single ligand has been reported [13].

ACCEPTED MANUSCRIPT

1.4 EGFR and ErbB receptor homo- and heteromeric interactions

The formation of dimers or higher oligomers is essential for the activation of all receptor protein tyrosine kinases including the ErbB family. Members of the ErbB family are capable of forming homodimers and heterodimers in combination with each other. The dimerization arms in the extracellular domain II share great homology between all members of the family [14]. As an example, the EGFR is able to form heterodimers with ErbB2, ErbB3 and ErbB4 after stimulation with the natural ligand EGF [15]. The ErbB2 (Her2) receptor is not able to bind any physiological known ligand and therefore the formation of a homodimer is unlikely [3]. The ectodomain of ErbB2 is fixed in an open conformation similar to a ligand-activated state, but unique structural features block the site where binding occurs in the other receptors of the family [7]. Since ErbB2 is always ready for dimerization, it is the most favorable binding partner for the other ErbB family members. Overexpression of ErbB2 receptor, which is often observed in the pathogenesis of different types of cancer, can lead to the formation of active homodimers [3], thus stimulating cell signaling processes. In contrast to the other family members, the ErbB3 receptor lacks a catalytically active kinase domain, and homodimerization cannot activate the tyrosine kinase signaling. Nevertheless, ErbB3 is able to activate the kinase domains of the other ErbB family members. Surprisingly, the heterodimer formed by ErbB2 and ErbB3 is known to build the most robust signaling complex in the ErbB family [3].

2 Intracellular protein-protein interactions

Upon ectodomain dimerization or activation respectively, the transmembrane region transfers the signal to the intracellular domain and initiates the formation of an asymmetric dimer of the two kinase domains, which starts the downstream signaling. One kinase domain therein acts as an activator domain, whereas the other domain acts as the receiver domain (Figure 1) [16]. Reasons why structural similar EGFR ligands can induce various biological effects was studied over the past decades [17]. Based on molecular dynamic calculations the presence of two different

ACCEPTED MANUSCRIPT

transmembrane helix states was proposed: an N-terminally associated and a C-terminally associated transmembrane dimer. Simulations by Arkhipov et al. showed that the N-terminal dimer corresponds to the active EGFR dimer, whereas the C-terminal transmembrane dimer is associated with the inactive EGFR dimer. This hypothesis is supported by mutational analyses [18]. On the intracellular side, the juxtamembrane region (residues 645-682) [19], which is located in direct proximity to the transmembrane region, attained attention. This region can be subdivided into two segments: the N-terminal half (JM-A) and the C-terminal half (JM-B). It was shown that the JM-A segment forms an α-helix and that upon dimerization these helices interact in an antiparallel manner [19]. Research suggested that ligand binding sterically influences these helices, which could explain why different EGF ligands induce different cellular responses. Doerner et al. found three different JM-A conformations upon binding of different natural ligands [20]. These different conformations might enable recruitment of different proteins known to bind the JM-A region and might be responsible for various different biological signaling pathways [20].

2.1 ErbB signaling network and associated proteins

Several of the main intracellular protein-protein interaction pathways are involved in the downstream signaling processes of the ErbB receptors. These include the phosphatidylinositol 3-kinase (PI3K)/AKT signal route, the Ras-MAP kinase pathway and the phospholipase C (PLCγ) signaling. Taking all processes into consideration the output of the ErbB signaling is ranging from cell division and migration to adhesion, differentiation and apoptosis. The specific output depends on the ErbB dimer pair and the specific ligand [21]. The PI3K/AKT signaling route is involved in cell survival. The PI3K is recruited by the phosphotyrosines of the ErbB receptor and phosphorylates the membrane-bound substrate phosphatidylinositol-4,5-bisphosphate (PIP_2) into phosphatidylinositol-3,4,5-trisphosphate (PIP_3). Because of its high affinity towards PIP_3 the protein kinase B, also known as AKT, is recruited to the membrane following the activation of mTOR (mammalian target of rapamycin), which is a key component of many cellular processes. [2,22,23]. For activation of the Ras-MAP kinase pathway, autophosphorylations of the receptors' C-terminal tyrosines result in

ACCEPTED MANUSCRIPT

docking sites for several proteins, such as Grb2. The adaptor protein Grb2 itself serves as a docking site for other proteins, e.g. the Ras-guanine nucleotide exchange factor SOS (son of sevenless). SOS then activates the Ras/MAP kinase pathway which leads to the activation of ERK. ERK1/2 has several substrates including transcriptions factors that are involved in cell division processes [2]. The third important EGFR-activated pathway is the PLCγ cascade. This enzyme is able to bind to the PLCγ phosphotyrosine binding sites and upon activation catalyzes the hydrolysis of PIP_2 into inositol-1,4,5-trisphosphate ($IP_{3)}$ and diacyl-glycerol (DAG). IP_3 is an upregulator of the calcium ion concentration in the cell, whereas DAG activates the protein-serine/threonine kinase C (PKC). The PKC acts on several substrates and can thus induce several physiological processes [2,24]. A brief overview of the signaling pathways of the EGFR is given in Figure 2.

2.2 EGFR downregulation and internalization

Ligand binding induces a series of events in the downstream machinery of the EGFR and a strong downregulation is required in order to prevent abnormal cell growth. For the control of the EGFR response, the receptor is clustered over clathrin-coated regions on the cell membrane. After internalization within clathrin-coated vesicles, the vesicles fuse with early endosomes [21]. Inside these early endosomes recycling of free EGFRs to the membrane is possible whereas ligand-bound receptors undergo a degradation process. During maturation of the endosome the C-terminal domain of the receptor is segregated to prevent interaction with proteins of the signaling cascade. Late endosomes fuse with lysosomes that break down the ligand-receptor complexes by lysosomal proteases [25]. A second route of downregulation and degradation of the EGFR is the ligand-induced polyubiquitination, which is mediated by the intracellular adapter protein Cbl. Cbl acts as an E3 ubiquitin-protein ligase and binds via SH2 proteins to a specific tyrosine residue in the C-terminal region of the EGFR [26]. The RING finger domain of Cbl recruits the E2 ubiquitin-conjugating enzymes. The polyubiquitination of the receptors leads to degradation of the receptor by the 26S proteasome [25]. Both downregulation processes play an important part for the signaling and aberrant regulation can promote cancer.

ACCEPTED MANUSCRIPT

3 Extracellular protein-protein interactions

Extracellular interactions between proteins and the EGFR can be differentiated into interactions of the receptor with its natural ligands and with targeted therapeutic agents.

3.1 EGFR natural ligands

Seven different peptide growth factors are known to bind and activate the EGFR. The EGF-family includes EGF, transforming growth factor α (TGFα), heparin-binding EGF-like growth factor (HBEGF), amphiregulin (AREG), betacellulin (BTC), epiregulin (EREG) and epigen (EPGN) [27]. All peptides are synthesized as type I transmembrane proteins with a cytosolic C-terminus and an extracellular N-terminus. The peptide growth hormones are released from the membrane by an enzymatic cleavage via proteases called ADAM (a disintegrin and metalloprotease). ADAMs belong to the class of membrane-anchored metalloproteases and play the most important part in protein ectodomain shedding [28,29]. Crystallographic studies of the EGF:EGFR interaction by Ogiso et al. revealed that the EGF molecule interacts with three different interaction sites named 1 to 3 [5]. The main interactions are illustrated in Figure 3A. Site 1 is located on domain I of the receptor and is in contact with the B loop of EGF, which forms a β-parallel sheet (residues 20-31). The contact is characterized by hydrophobic interactions between Leu14, Tyr45, Leu69 and Leu98 of the EGFR and Met21, Ile23 and Leu26 of EGF. The other two interactions sites are located in domain III of the receptor (Figure 3B). The A loop region (residues 6-19) interacts with site 2 and the C-terminal region with site 3. Site 2 interactions mainly consist of hydrophobic interactions between the receptor amino acids Val350 and Phe357 with Leu15 and Tyr13 of the ligand. Furthermore, a salt bridge between receptor Asp355 and ligand Arg41 enhances the strength of interaction. The contact area of site 3 is characterized by hydrophobic interaction between Leu382, Phe412 and Ile438 of the receptor and Leu47 of EGF as well as hydrogen bonds between the Gln384 side chain of the EGFR and Gln43 and Arg45 (EGF) [5]. The precise binding epitopes for the other ligands are not known, as there are no crystallographic data available.

ACCEPTED MANUSCRIPT

3.2 Recent progress in therapeutically EGFR targeting

The EGF receptor and its ligands are an important regulatory system to control the cell state and differentiation. Dysregulation of these pathways can lead to several diseases such as cancer, diabetes and autoimmune, cardiovascular, inflammatory and nervous disorders [30]. Growth of human malignant tumors is often driven by the overexpression or constitutive activation of ErbB receptors. Analyses of patient samples demonstrated EGFR overexpression in several types of cancer, which in several cases was associated with poor prognosis [31]. Unfortunately, a common definition of overexpression and a widely accepted standard for determining the EGFR status remains elusive [32]. So far three antibodies targeting the EGFR obtained regulatory approval: Cetuximab for EGFR-expressing metastatic colorectal cancer and squamous cell carcinoma of the head and neck (SCCHN), Panitumumab for metastatic cancer of the colon or rectum with wild-type Ras, and Necitumumab for advanced squamous non-small-cell lung carcinoma (NSCLC) [33].

Targeting of the EGFR and blocking the ligand binding site was proposed already in the early 1980s [34]. First experiments in cell culture proved that competitive binding of a monoclonal antibody to the receptor can disrupt EGFR signaling and hinder cellular proliferation [34]. Since then several efforts focused on the development of EGFR antagonists on the basis of monoclonal antibodies (mAbs) directed towards the ectodomain. In contrast small molecule inhibitors such as IRESSA (Gefitinib) and TARCEVA (Erlotinib), which are not in the focus of this review, bind to the intracellular catalytic domain of the tyrosine kinase and inhibit autophosphorylation.

3.2.1 Antibodies

Monoclonal antibodies (mAb) typically had been selected for blocking EGFR activation. This is exemplified by the monoclonal antibody 225, which advanced to the clinic as chimeric C225 with the international nonproprietary name (INN) Cetuximab and which is marketed as Erbitux. The monoclonal antibody C225 was one of the first biological EGFR antagonists brought to the clinic and the first EGFR directed mAb approved by the US Food and Drug Administration (FDA) and the

ACCEPTED MANUSCRIPT

European Medicines Agency (EMA) for treatment of colorectal and head and neck cancer in combination with chemotherapy in 2004. C225 binds to site 3 of domain III (residue 449-463) of the EGFR whereby the light chain of the antibody interacts with the C-terminal region of the binding site. C225 binding prevents EGF binding and sterically hinders the adoption of the extended conformation required for dimerization [35] (Figure 4A). Upon binding, Cetuximab also initiates internalization and subsequent degradation of the receptor resulting in a block of the cell cycle, inhibition of proliferation and induction of programmed cell death [36].

A further mAb co-crystal structure with domain III is available for Matuzumab (EMD 72000), which has been in several clinical trials [37]. Matuzumab was derived from the mouse monoclonal antibody mAb425. This antibody binds in proximity to Cetuximab and the natural ligands, but with little to no overlap on the β-helix of domain III (Figure 4A). The main interaction with the receptor is characterized by about 11 hydrogen bonds and two salt bridges that stabilize the binding. The crystallographic structure of the Fab fragments of Cetuximab and Matuzumab are shown in Figure 4A in combination with domain III of the EGFR. The binding epitopes for the natural ligand EGF and the Fab fragments of both antibodies are mapped on the surface of domain III of the EGFR in Figure 4C-E.

Panitumumab (E7.6.3, ABX-EGF), as the first fully human mAb and active ingredient of Vectibix, was approved in 2006 for metastatic colorectal cancer. This antibody was isolated from antibody-producing XenoMouse strains that carry the genetic information for the human antibody gene repertoire [38]. After immunization of those mice with A431 epidermoid carcinoma cells, that are known to overexpress the EGFR [39], hybridomas were screened for EGFR-binding antibodies [38], resulting in a mAb that binds to EGFR with high affinity, blocks the receptor and disrupts downstream signaling [40].

The latest antibody approved is Necitumumab (IMC-11F8, Portrazza), which was approved in late 2015 to treat metastatic squamous non-small cell lung cancer (NSCLC). The Fab fragment of Necitumumab was found by screening a naïve human Fab phage display library for high affinity

ACCEPTED MANUSCRIPT

binders to A431 cells even in competition with Cetuximab. The crystallographic analysis of the binding mode showed that Necitumumab hits exactly the same epitope as Cetuximab. However, the mode of binding differs between both antibodies [41].

Monoclonal antibodies of the IgG class have a molecular weight of about 150 kDa. Size, weight and composition of these molecules greatly influence the biodistribution and the ability to penetrate a solid tumor [42,43]. Tumor penetration of large molecules can be restricted and therefore multiple ways to improve this ability are tested. One typical approach is to reduce the size of the targeting molecule. To this end different antibody and antibody-mimetic formats were engineered: antigen-binding fragments (Fab), single-chain variable fragments (scFv), nanobodies (camel V_H) or domain antibodies (camelized V_H) as well as "third generation" molecules. These molecules can then be multimerized for example by attaching multimerization domains to generate miniantibodies [44] or by daisy chaining the modules. Fab fragments represent the oldest class of antibody fragments. Historically, they were created by proteolysis of a monoclonal antibody but nowadays they can be produced directly in eukaryotic or prokaryotic expression hosts. Up to now three different Fab fragments were approved for therapeutically uses by the FDA, however none of them for therapy of cancer [45–47]. Single chain variable fragments (scFvs) are recombinant molecules that are comprising the variable regions of light and heavy chains of the antigen-binding region are joined by a flexible linker, which is designed to maintain binding affinity and stability [48]. scFvs have size of about 27 kDa and therefore renal clearance from the circulation is rapid [49] leading to an overall efficiency that is still lower than desired for therapeutic approaches. This limitation can be overcome by increasing the size for example by dimerization and multimerization of scFvs [48].

To enhance efficacy of EGFR targeting, research and development increasingly focuses on the development of novel binding molecules or novel binding modes such as bispecific antibodies. Bi- or multispecific targeting of a molecule enables to reach cells expressing different cancer specific antigens and avoid escape or alternatively can direct immunological cytotoxicity to tumor cells. Early work demonstrated that EGFR targeting scFvs based on the 425 antibody can be combined with

ACCEPTED MANUSCRIPT

further specificities via heteromerization modules [50,51]. In the BiTE (Bi-specific T-cell Engager) format scFvs are connected via another linker, joining a T-cell engaging anti-CD3 scFv and a tumor-specific (for example anti-EGFR) scFv fragment [52]. A new approach is a single antibody variable domain that recognizes more than one antigen. This two-in-one or dual-specific antibody concept was demonstrated by Bostrom et al by mutating the light chain complementary determining regions (CDRs) of a monospecific Her2 antibody mainly interacting via the heavy chain CDRs such that also VEGFR can be bound via the light chain [53]. This technique was also used to generate a bispecific antibody directed against EGFR and HER3 [54] by modifying a monospecific EGFR antibody that was selected from a phage display Fab library with diversity restricted to the heavy chain. A single, high affinity clone was then further diversified with a library of mutations in the light chain CDRs and this library was selected for binding to both receptors. The two-in-one antibody Fab fragment named DL11f (or MEHD7945A when reformatted as IgG) was characterized by crystallization and its epitope on domain III overlaps with that of Cetuximab (Figure 4B) [54].

A further approach to modulate efficacy and mode of action of antibodies is the engineering of the Fc-domain. Based on the knowledge of the biological effector functions of natural antibody isotype variations the engagement of the Fc-receptors and thus complement-dependent cytotoxicity (CDC), antibody-dependent cell mediated cytotoxicity (ADCC), and antibody dependent cellular phagocytosis (ADCP) can be exploited. Immune cells are recruited by Fc receptors and since most therapeutic antibodies belong to the IgG class FcγRI, FcγRIIa and c, and FcγRIIIa and b are typical targets [55]. Testing the genetic *FCGR* status can predict clinical outcome as seen in Cetuximab treated colorectal cancer patients [56]. In addition, half-life is influenced by binding to the neonatal Fc receptor FcγRn and this interaction can be optimized as well [42].

Another approach of increasing the potency and efficacy of monoclonal antibodies is the generation of antibody-drug conjugates (ADCs). Monoclonal antibodies are often lacking the ability to trigger long lasting therapeutic effects. An enhancement of the activity can be reached by coupling the monoclonal antibody to a potent drug by a cleavable linker. The antibody is able to target cells

ACCEPTED MANUSCRIPT

expressing specific antigens. After internalization via endocytosis, the ADC is degraded and the cytostatic agent is released and can induce apoptosis [57,58]. The first EGFR antagonist to be in clinical trials is ABT-414, which is built up from a human IgG linked with monomethyl auristatin F (MMAF). This anti-EGFR ADC is currently tested in phase I/II for the treatment of squamous cell tumors [59].

Last but not least antibodies targeting the EGFR can be combined for enhanced efficacy. Sym004 and MM-151 are formulations of two or three monoclonal antibodies, respectively [60,61]. The latter was designed to overcome high affinity ligand signaling and was also shown to overcome therapeutic evasion by extracellular domain mutations [62].

3.2.2 Repeat proteins

In nature, binding of proteins is often a feature of repeat proteins. These proteins are characterized by homologous structural motifs, which stack together in elongated structures. Homologous structural motifs have a size of about 20 to 40 amino acids that can range from simple to complex structures [63]. Two different repeat proteins for EGFR protein-protein interaction are described in the following.

Ankyrin repeat proteins were first found in the human erythrocyte protein ankyrin. The L-shaped subunit structure is composed of usually 33 amino acid residues forming two antiparallel α-helices followed by a β-turn [64]. They bind to other proteins via the extended groove formed by the repetitive structure. Engineering of such ankyrin repeat proteins resulted in so called designed ankyrin repeat proteins (DARPins) [65]. Sandwiched between capping domains at the N- and C-terminal end, which enhance stability and enable bacterial expression, the structure presents a binding surface to the environment [66]. Binding to various targets of interest is possible through residue randomization combined with ribosome or phage display selection. Phage display was used to select four EGFR binding DARPins named E01, E67, E68 and E69 [67]. The binding epitopes were mapped using alanine scanning mutation experiments that highlighted important amino acid

ACCEPTED MANUSCRIPT

residues for binding (Fig 5A-B)[68]. Three DARPins (E01, E67, E68) bind to domain III and one to domain I (E69) of EGFR. To further improve efficacy, two DARPins with different epitopes (E01 and E69) were linked either by a flexible linker or by a dimerizing leucine zipper domain, while the latter results finally in a tetrameric construct with two specificities. The tetravalent construct induced G1 arrest in A431 cells and inhibited receptor recycling [69].

A second repeat protein studied for targeting is a structure called 'Repebody'. Their basic structure was derived from leucine-rich repeats (LRR). LRRs are built up from short repeats of about 20 amino acids that are organized in a β-strand-turn-α helix structure. They are present in more than 2000 proteins that mediate a broad range of protein-protein interactions [70]. In variable lymphocyte receptors (VLRs) they play an important role in the adaptive immune system of jawless vertebrates and are therefore able to act as immunoglobulin substitutes [71]. The LRR structure was also used to design a scaffold for targeted interactions, which was named repebody. In a current study this structure was directed against the EGFR and additionally a cytostatic drug was attached [72]. The interaction between the repebody and the EGFR domain III was shown in a crystal structure analysis (Figure 5C). Interaction occurs with a binding affinity in the sub-nanomolar range in domain III overlapping with the epitope of the antibody Cetuximab [72].

3.2.3 Affibody

For identification of tumor metastases in the body radionuclide molecular imaging is a standard method albeit with limited availability of tumor specific targeting molecules for radionuclides. Since these agents can conceptually also be used for therapy they are often referred to as theranostics. Utilizing the ErbB family receptors for imaging has been tested with radiolabeled monoclonal antibodies and radiolabeled natural ligand EGF. Tests with antibodies demonstrated the general applicability but using mAbs such as Cetuximab can be hampered by poor contrast due to their long presence in the circulation and slow accumulation in the specific tumor tissue [73]. Radiolabeled EGF can yield improved tumor to blood rations due to its smaller size and the resulting quick renal

ACCEPTED MANUSCRIPT

clearance and improved malignant tissue penetration. However, EGF based imaging needs higher concentrations and can induce also adverse reactions possibly due to agonistic properties [74]. For imaging, targeting molecules are often labeled with DOTA which chelates radioisotope such as ^{68}Ga, ^{86}Y or ^{55C}Co for the use in PET (positron emission tomography) scans or ^{111}In for SPECT (single-photon emission computed tomography) scans. The developers of a small binding molecule named Affibody aim at combining small size with high affinity for imaging in a molecule that can be produced in bacteria of by chemical synthesis [75]. Affibodies are derived from the Z-domain of staphylococcus protein A and have a size of about 6.5 kDa. The 58 amino acid scaffold forms a three helix bundle where 13 amino acids of helix 1 and 2 are randomized in order to increase affinity [76,77]. Site specific conjugation of DOTA via a cysteine to an EGFR targeted Affibody (DOTA-$Z_{EGFR:2377}$) was evaluated in mice xenograft models. The bio-distribution showed that a radiolabeled Affibody can be used to visualize malignant tumors [78].

3.2.4 Fibronectin/Monobody/Adnectin

Monobodies, also proprietarily named Adnectins, represent a further family of designed targeting proteins. They are based on the 10th type III domain of fibronectin (^{10}Fn3). Fibronectin is an abundant human protein of the extracellular matrix, which is based on a modular structure. The structure of ^{10}Fn3 and the variable fragment of an antibody resemble a similar topology even if they lack any sequence homology. Like the variable region of antibodies, the ^{10}Fn3 domain is built up from a sandwich of two anti-parallel β-sheets with loops at each pole of the molecule [79–81]. Unlike the antibody domain, the ^{10}Fn3 domain has no disulfide bridges or free cysteine residues. Using the ^{10}Fn3 domain as a starting point, Fibronectin monobodies were designed that have target-binding properties [81] and subsequently EGFR targeted monobodies were selected by yeast surface display (e.g. EI3.4.3) [82] or mRNA display (e.g. E#1, Adnectin 1) [83,84].

Adnectins are suitable for multimerization, and bispecific Adnectins can be generated by connecting two moieties with a flexible glycine-serine linker sequence. Such a bispecific molecule with a ten amino acid linker, a specificity for EGFR and for IGF-IR as well as a PEG-extension (EI-tandem),

ACCEPTED MANUSCRIPT

showed improved pharmacological properties [83]. Fibronectin based targeting proteins can also be used for radionuclide molecular imaging upon chemical coupling with DOTA. An EGFR targeted DOTA monobody complexed with ^{64}Cu was stable in serum for at least 24 h and exhibited good tumor localization in tumorgrafted mice. The only disadvantage was the high renal uptake and retention which is often a problem for small antibody mimetics [85]. Protein-protein interaction between Adnectin1 and the EGFR extracellular domain was analyzed by crystallography [86]. Co-crystal structures revealed that Adnectin1 is interacting with domain I of the EGFR. The epitope for Adnectin1 is presented in Figure 6A in the context of the whole extracellular domain of the EGFR and in detail in Figure 6B.

3.2.5 VHH/Nanobodies

In the early 1990s a new type of antibodies was discovered in the camelidae family (dromedary, camel and llama). In contrast to the well characterized human and rodent antibodies, which consists of heavy and light chains, the camelid antibodies naturally are only formed of heavy chain homodimers and they also lack the second constant domain. These antibodies, which are also referred to as heavy-chain antibodies (HCAb), have only one single immunoglobulin domain (Ig) variable fragment (VHH) with three CDRs forming the antigen-binding domain, which is structurally slightly distinct from the artificially generated heavy chain Fv domain antibodies. The three camelid CDRs are in average larger in size compared to human antibodies, which presumably compensates the lack of light chain CDRs and allows for a higher flexibility in antigen binding [87]. For the VHH domain the proprietary name 'Nanobody' has been introduced. These small fully functional antigen-binding fragment have a size of 15 kDa. They are typically obtained from immunized llama VHH libraries, which are recloned and selected by phage display [87]. Well behaved nanobodies are easy to produce in bacteria and they do not require artificial linkers as in the case of the scFv [88]. As with the other antibody mimetics, the therapeutic approaches with VHH antibody fragments can be divided in three different categories based on their application. First, naked and multimerized

ACCEPTED MANUSCRIPT

nanobodies are used as antagonists. Second, nanobodies are fused to effector domains, and third nanobodies are presented on the surface of nanoparticles such as liposomes [89].

Nanobodies specifically targeting the EGFR were selected and optimized using phase-display and studied for their cellular and structural properties [90–92]. The nanobodies 7D12, EgA1 and 9G8 were compared to existing antibodies Cetuximab and Matuzumab as well as to EGF with regard to their binding site by Schmitz et al 2013 (Figure 7) [91]. The co-crystal structure of VHH 7D12 with domain III of EGFR shows that 7D12 interacts with domain III of the receptor primarily on the EGF binding site 2 and thus overlaps only partially with the EGF epitope, since EGF binds to site 2 and site 3 [5] (Figure 7A-C). The interactions of 7D12 with domain III is mainly mediated by salt bridges and hydrogens bonds between amino acid residues Asp355, Arg353, Phe357 and Gln384 of domain III and Arg30, Glu100f and Asp101 of the nanobody, respectively. Some of these amino acid interactions are also important for binding of the natural ligand EGF such as e.g. Asp355 and Phe357. The binding site of EgA1 in comparison to the antibody Matuzumab site is shown in Figure 7D-F. Both epitopes are on the same side but show only minimal overlap.

A bispecific construct, in context also called biparatopic, which combines two VHH fragments with two distinct binding sites on EGFR and a linker which is too short to allow for binding to a single receptor was constructed. In cellular studies, this construct significantly increased receptor internalization and in combination with juxtamembrane mutations a model of receptor clustering and internalization was inferred [93]. An even higher order of multivalency was obtained with nanoparticles, which were used for drug delivery [94,95].

3.2.6 Peptides

Peptides have the advantage of being able to diffuse in to tissue and having a low immunogenicity, yet they are prone to degradation and rapid clearance. Thus, peptides are often used in a stabilized form or as a modular targeting domain in a larger context. Several approaches have been used to identify EGFR targeted peptides. EGF itself, which is 53 amino acids long and stabilized by three

ACCEPTED MANUSCRIPT

disulfide bridges, was the first choice to deliver e.g. cytostatic drugs into cancer cells. However, the mitogenic and neoangiogenic activity hampered further clinical development.

A novel targeting peptide was identified using a commercial peptide phage-display library and the extracellular domain of EGFR as target. The 12 residue peptide GE11 (sequence: YHWYGYTPQNVI) was shown to have no mitogenic activity and it was successfully internalized by EGFR overexpressing cells. Surface plasmon resonance experiments measured an approximate K_d of 22 nM. The fact that the affinity is lower compared to EGF might be explained by the much smaller size of GE11 and the structural flexibility [96]. Peptide GE11 was used in several further studies.

Performing phage display against an EGF-binding mAb resulted in a peptide (sequence: EHGAMEI) that should mimic EGF. Analysis of this peptide indeed showed antagonistic behavior [97].

A further peptide that was characterized is D4 (sequence: LARLLT), which was identified by a computer-assisted design process using the available EGFR crystal structures. A surface pocket in EGFR domain I served as a starting point, which is not located in or near the epitope of the natural ligand EGF. Six amino acid residue peptides were computationally sampled and finally 10 different six residue peptides were obtained. Peptide D4 showed significant binding to the receptor in cell culture experiments. A liposome was decorated with the D4 peptide and experiments showed an uptake of the particle. Also, tumor-specific accumulation was observed in a xenograft model [98].

Another starting point for the design of peptide EGFR antagonists was the EGFR itself. The dimerization of the extracellular domain is mainly depended on interaction of the dimerization arm in domain II of the receptor. A peptide mimicking the EGFR dimerization arm should thus interact with the receptor. Mizuguchi et al. analyzed various disulfide clamped version of the EGFR dimerization arm sequence [99]. Hanold et al. generated a peptidomimetics of the dimerization arm named EDA (e.g. EDA1: *Aha*YNPTTYQ*Pg*) in which the termini are covalently connected via a triazolyl bridge or in a further study by selenylsulfide-bridges [100]. It was shown that this proteolytically stable peptide is able to downregulate dimerization, phosphorylation and downstream signaling. Due

ACCEPTED MANUSCRIPT

to the small size peptides are an easy to use fusion partner and are well suited for targeting approaches involving chemical methods.

Also intracellular peptides for EGFR inhibition were identified. A six amino acid peptide named Disruptin represents a cytosolic wild-type EGFR segment that is important for dimerization and interacts with the heat shock protein 90 (HSP90.) The peptide infers with dimerization and promotes degradation [101].

3.2.7 Nanoparticles and viral particles

An emerging field of EGFR targeting is the delivery of larger entities such as viral particles and nanoparticles. In 'Trojan horse' strategies they allow for the delivery of lager amounts of payloads or more complex payloads such as gene assemblies [102]. Nanoparticles have a size between a few to several hundreds of nanometers, a range equivalent to that of larger proteins to cellular organelles [103]. Nanoparticles offer opportunities for the delivery of cytotoxic agents, certain physical effectors or agents for molecular imaging. They can be manufactured from inorganic compounds or by means of organic chemistry as polymer structures or as biological or biomimetic structures such as liposomes or protein polymers. The surface of these shells is functionalized for targeting and the cargo can be embedded within the particle or be attached to the surface as well. Nanoparticles can be functionalized with different structures, e.g. antibody fragments or peptides, and provide opportunities for multi-targeting. For example, albumin nanoparticles were targeted with an anti-EGFR nanobody and loaded with a kinase inhibitor [104]. Li et al. used maleimide-functionalized poly(ε-caprolactone) bovine serum albumin particles coated with Cetuximab and additional ^{131}I arming for analyzing effects on various cell lines [105]. A Fab C225 fragment or an anti-EGFR scFv were used to specifically deliver liposomes to tumor cells [106]. Also, the EGFR binding peptide D4 directed liposomes to tumor cells *in vitro* and *in vivo* [98].

For the delivery of genes viruses or viral-like particle (VLPs) provide sophisticated designs. Replication competent oncolytic viruses address tumor cells mainly based on their physiology whereby virus-like

ACCEPTED MANUSCRIPT

particles are typically addressed via capsid modifications. In VLPs, targeting is preferably introduced by genetic means during viral particle production albeit chemical coupling similar to that used for nano-particles is also an option. The delivery of genes can be used to compensate cancerous mutations and rehabilitate normal cell function or provide the basis for what is referred to as molecular chemotherapy. Virus directed enzyme prodrug therapy (VDEPT) is based on the introduction of an enzyme that can activate a prodrug in cells. The activation leads to a cellular suicide and, when taking advantage of the bystander effect, also eradication of neighboring cells. This approach was shown to work in cell culture for the non-enveloped human adeno-associated virus serotype 2 (AAV2) by Hagen et al [107]. In this work, an EGFR targeting DARPin (E01) or Affibody ($Z_{EGFR:1907}$) were genetically fused to a capsid protein, and upon EGFR-dependent transduction of target cells enzymatic conversion of either the prodrug 5-fluorouracil or ganciclovir lead to apoptosis. An enveloped oncolytic measles virus was retargeted to EGFR or EGFRVIII using the scFv 425 or MR1, respectively, fused to protein H6. The enveloped oncolytic herpes virus HSV1716 was targeted to EGFR expressing cells via a scFv (derived from mAb108) fusion to surface protein gD [108]. Adenovirus was retargeted to EGFRVIII by including the peptide PECPHC1 [109]. A second approach, which may require less genetic engineering of the viral production, is the use of adaptor proteins that bind the virus and the target. Verheije et al. used the nonhuman coronavirus for targeting the EGFR by using adaptor proteins, which they encoded in the viral genome. These proteins consist of a soluble domain that targets coronavirus natural receptor fused to EGF (mCEACAM1a EGF fusion). Experiments showed that these viruses are active as oncolytic agents *in vivo* [110].

4 Expert commentary

Several proteins and peptides are available that target different parts of the EGFR and a plethora of data from detailed structures to cellular effects and in some cases even for clinical performance are accessible for these molecules. Modular combinations of these targeting molecules have been

ACCEPTED MANUSCRIPT

studied and binding molecules have been tested for the delivery of linked compounds ranging from radioactive isotopes via chemicals and toxins to viral- and nano-particles.

Despite all these successes targeting of EGFR is still not an off the shelf endeavor particular when the goal is a clinical application. EGFR overexpression and signaling is a moving target. There are high and low affinity-, auto tethered-, extended-, preclustered-, homo-, and heteromeric forms of EGFR which can be of wild-type sequence or mutated and which can be connected to several major downstream signaling pathways, and ultimately tumors might evolve resistance by introducing additional alterations in the receptor itself or the downstream signaling.

Due to this complicated context, in some cases experimental reporting could be improved. Even in seemingly stringently controlled cellular experiments quantitative data such as internalization may be confounded by unreported experimental details such as previous cell culture conditions. In addition, terms such as “rapid internalization” need to be provided as biophysical time constants and/or in the context with a reference system such as cetuximab.

Many clinically relevant engineering ideas have been developed, but clinical testing of all variations is neither ethically nor financially justified. Unfortunately, simple mouse xenograft models are of limited predictive value for clinical success and the lack of predictive and manageable preclinical models for the various EGFR overexpressing tumor types poses a current bottleneck. For some targeting approaches, in particular the smaller sized antibody mimetics, the theranostic approach might help to first obtain pharmacokinetic and distribution data before moving to the more demanding therapy.

As seen by the rise of data for heavily engineered antibodies and antibody mimetics, future therapeutic molecules will have to address more than just binding.

ACCEPTED MANUSCRIPT

5 Five-year view

Within the next five years the now available cellular data on receptor activation and internalization as well as the clinical data on EGFR therapy evasion will guide further design. Rather than simply blocking EGFR activation new therapies will continue to harness additional modes of action as there are bispecific targeting, antibody drug conjugates, additional immune system recruitment. Using EGFR guided viral particles will enable to impose further levels of specificity at the transcriptional and translational level while at the same time exploiting the multivalency of surface modifications. As binding scaffold of choice antibodies will still dominate but antibody mimetics will gain ground due to lower manufacturing costs in particular for multispecific constructs. For bringing drugs or genes into the cell, internalization without activation is desired, for immune system recruitment binding without internalization is desired, and in both cases concomitant blocking of EGFR activity should occur and healthy cells with low amounts of EGFR should be spared. The best candidates to achieve these goals are hard to predict from the current data and likely there is more than one good solution.

6 Key issues

- Epidermal growth factor receptor (EGFR) is overexpressed in various types of cancers and can drive malignancy. It is a validated marker for targeted cancer therapeutics either by disrupting its signaling activity or by using it as a surface marker for cancerous cells.
- EGFR targeting by biotherapeutics is almost exclusively achieved by binding the extracellular domain. Peptides and the following classes of proteins have been developed for EGFR targeting: antibodies, antibody fragments, cameloid VHH, DARPin, Affibody, Fibronectin Domain Fn10. Many EGFR targeting proteins and peptides are selected to compete for EGF binding and bind to domain III of EGFR, but binders for others sites are available.
- EGFR signaling is well studied, but the molecular mechanisms of how the various external binding events are mechanistically coupled to signaling and internalization are still under

ACCEPTED MANUSCRIPT

investigation. Biotherapeutic blocking of EGFR signaling is achieved by interfering with ligand binding, blocking dimerization, or induction of downregulation or by inhibition of the kinase activity with small molecules.

- Using EGFR as cellular marker requires discrimination from healthy cells, which express lower amounts of EGFR but are more abundant. Therapies need take into account that cancers are heterogeneous and can evolve to sidestep EGFR overexpression or signal dependence
- Researchers responded to the complex requirements of EGFR targeted therapy, by developing multifunctional molecules beyond antibody binding and natural ADC and ADCC.

Funding

This paper was not funded

Declaration of interests

The authors have no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

ACCEPTED MANUSCRIPT

Figure 1: Schematic representation of the EGFR structure as a monomer and as a dimer. The N-terminal extracellular domain is build up from four domains I-IV followed by a transmembrane helix, a juxtamembrane segment and the kinase domain. (A) As a monomer without ligand the receptor is mainly present in the autoinhibited form in which domain II and IV are tethered by the dimerization arm. (B) Natural ligands are bound in an open conformation sandwiched between domain I and III, whereby the dimerization arm is exposed promoting dimerization. After homo- or heterodimerization internalization starts and the intracellular kinase domains form an asymmetric complex initiating phosphorylation.

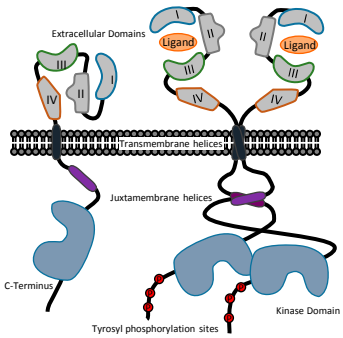
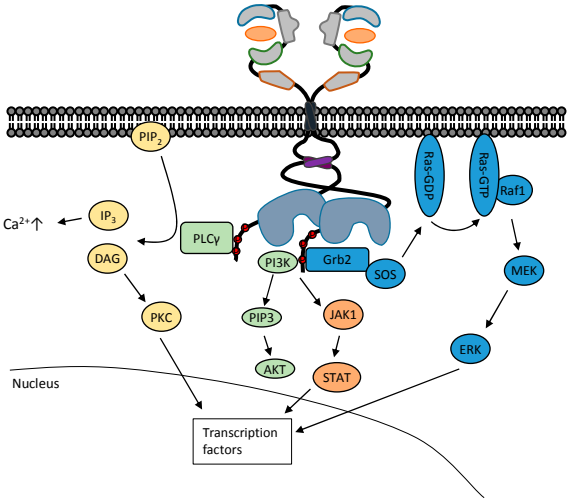
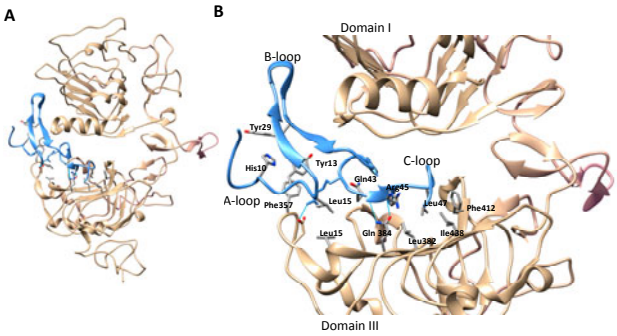


Figure 2: Schematic overview of the four main signaling pathways, that are activated upon receptor phosphorylation: the Ras-Map kinase pathway (blue), the JAK-STAT way (orange) as well as the PI3K (green) and PLCγ (yellow) pathways. All of these pathways are involved in processes controlling cell growth and proliferation.



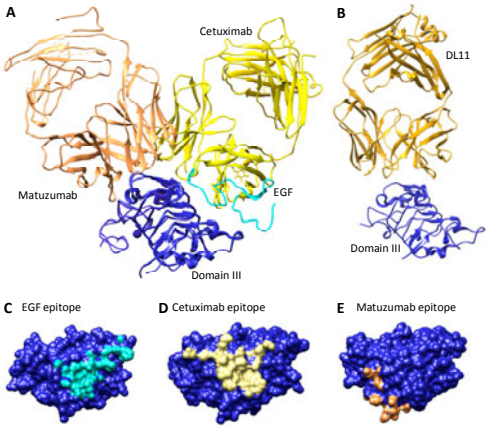
AC
VUSCRIPT

Figure 3: Crystal structure of an EGFR:EGF complex, illustrating the interacting sites. (A) Crystal structure of one EGFR ectodomain monomer (grey) in the open conformation bound to EGF (blue). (B) EGF binds to three different binding sites (1-3) on two different domains of the EGFR. The B-loop of the EGF (Cys14-31) interacts with site 1 located in domain I of the receptor. The A- and the C-loop are interacting with site 2 and 3, respectively. Amino acid residues are highlighted that mediate the main interactions between domain III and the EGF molecule. The binding is mainly stabilized by hydrophobic bonding between specific residues. Models were created from PDB [1IVO] with UCSF Chimera.



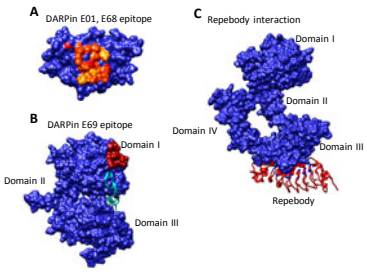
MANUSCRIPT

Figure 4: Crystal structure analysis of protein-protein interactions with domain III of the EGFR. (A) The position relative to domain III (blue) for the EGFR antagonists Matuzumab (brown) and Cetuximab (yellow) is shown in comparison to the natural ligand EGF (cyan). For reasons of simplification only the Fab fragments are shown in combination with domain III. (B) The dual specific Fab DL11 is shown in binding to domain III the same orientation as in (A). (C-E) The epitope of EGF (cyan), Cetuximab (yellow) and Matuzumab (brown) are mapped on domain III. Structures were created from PDB files: EGFR [1IVO], Matuzumab [3C08], Cetuximab [1YY9], and DL11 [3P0Y].



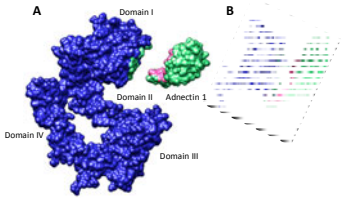
MANUSCRIPT

Figure 5: Crystal structure binding analysis of the EGFR targeted DARPin and repebody molecules. (A) DARPin E01 (red) and E68 (yellow) epitopes are mapped on domain III based on mutational analyses by Boersma et al.[68]. The merge area shows the overlap of both epitopes (orange). (B) DARPin E69 (red) binds to domain I of the EGFR as confirmed by mutation analysis experiments [68]. (C) The repebody interacts with domain III and is shown here in the context of an autotethered conformation. Structures were created from PDB files: EGFR [1IVO], repebody [4UIP].



MANUSCRIPT

Figure 6: Representation of the Adnectin1 binding. (A) Adnectin1 (green) with the interface region highlighted in pink interacts with domain I of EGFR (blue) with the interface highlighted in green. (B) For clarity the interaction is shown in a ribbon model of domain I and Adnectin1. The crystal structure was generated from PDB Adnectin1 [3QWQ].



ACCEPTED MANUSCRIPT

Figure 7: Binding mode of nanobodies to domain III of the EGFR. (A) Ribbon model of the nanobody 7D12 (pink) with domain III (blue). (B) The epitope of 7D12 is mapped on domain III in pink. (C) The epitope of the natural ligand EGF given for comparison in cyan on domain III depicted the same orientation. (D) The ribbon model of the EgA1 nanobody is shown interacting with domain III (blue). (E) The epitope of EgA1 is mapped on domain III in pink. (F) The epitope of Matuzumab is given for comparison on domain III in the same orientation. Structures were generated from PDB files: EGFR [1IVO], 7D12 [4KRL], EgA1 [4KRO], Matuzumab [3C08].

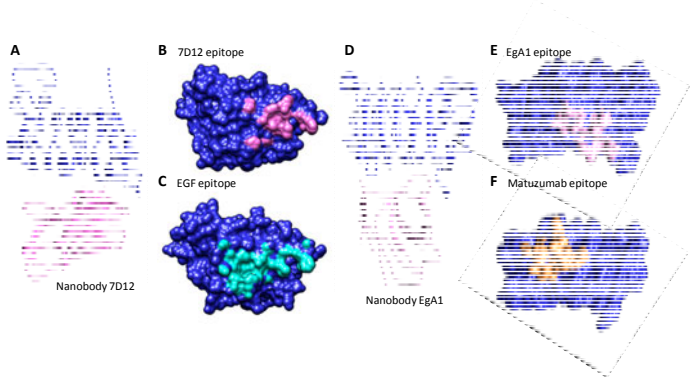


Table 1: List of EGFR interacting molecules

Molecule (Laboratory code, INN, Tradename)	Type of molecule; generation method; epitope	Rx[a]	Ref
Antibodies and antibody fragments			
IMC-C225 Cetuximab Erbitux	mAb IgG1, chimerized by Fv grafting; obtained from mice immunized with EGFR from A431; binds domain III	A	[111]
E7.6.3 ABX-EGF Panitumumab Vectibix	mAb IgG2; human; obtained from transgenic mice immunized with EGFR from A431; binds domain III	A	[40]
h-R3 Nimotuzumab TheraCIM	mAb; humanized by CDR grafting; obtained from mice immunized with placental EGFR; binds linear Epitope on domain III	(A)	[112]
IMC-11F8 Necitumumab Portrazza	mAb, humanized, obtained from a naïve Fab phage display library selected on A431 cells	A	[41]
2F8 Zalutumumab HuMax-EGFr	mAb, human, obtained from transgenic mice immunized with EGFR and A431; binds domain III	(A); D	[113]
425 EMD 55700	mAb; murine; mice immunized with EGFR from A431; used for radioisotope targeting	II; D	[114] [115]
EMD72000 Matuzumab	mAb; mAb 425 humanized by CDR grafting	II; D	[116]
ICR62	mAb, rat, obtained from rats immunized with EGFR from MDA-MB 468; binds domain III	I	[117]
Hum ICR62, GA201, RG7160 RO5083945, Imgatuzumab	mAb; humanized by CDR grafting; glycoengineered for bearing bisected, afucosylated carbohydrates	II; D	[118]
ch806	mAb; chimerized by Fv grafting; obtained from mice immunized with EGFRvIII (de2-7 EGFR) from fibroblasts; binds cryptic epitope on domain II	II	[119]
ABT-414	mAB ch806 monomethyl auristatin F conjugate	II	[120]
Sym004	mAb; mixture of mAB 992 and 1024; obtained from mice immunized with recombinant EGFR fragments; bind non-overlapping epitopes on domain III	II	[60]
MM-151	mAb, mixture of three human IgG1 monoclonal antibodies obtained by yeast display	I	[61]
D2C7	mAb; obtained from mice immunized with 14-mer peptide corresponding to the the EGFRvIII fusion junction; binds EGFR and EGFRvIII		[121]
D2C7-(scdsFv)-	mAb D2C7 fused with truncated Pseudomonas Exotoxin A		[12

ACCEPTED MANUSCRIPT

PE38KDEL			2]
13A9	mAb; blocks TGFα binding much stronger than EGF binding		[123,124]
29.1	mAB; obtained from mice immunized with paraformaldehyde-fixed A431 cells, binds carbohydrate on EGFR		[125]
scFv MR1	scFv; obtained from mouse immunized with the first 13 amino acids of the EGFRvIII N terminus and selection of a resulting scFv phage display library; used for toxin targeting		[126]
VHH 8B6, EGa1	VHH/nanobody; obtained from llama glama immunized with A431 preparations and phage display selection with EGF elution; 8B6 used for ^{99m}Tc targeting [127] and EGa1 for liposome targeting [95] [94]		[90]
VHH D10 and E10	VHH/nanobody; obtained from immunized llamas by alternating phage display selections on recombinant EGFR-human Fc fusion protein and A431 cells; bind only ligand free EGFR		[128]
VHH G10	VHH/nanobody; obtained from immunized llamas; binds EGFR with and without ligand; used for EGFR conformation studies		[128]
Designed binding molecules			
Affibody $Z_{EGFR:955}$; $Z_{EGFR:1907}$; $Z_{EGFR2377}$	Protein A Z-Domain scaffold; synthetic phage display library selected on EGFR-ECD with glycine–HCl elution, affinity maturation by directed evolution; ^{111}In-DOTA-$Z_{EGFR:2377}$ and ^{89}Zr-DFO-$Z_{EGFR:2377}$ used for imaging		[75,129]
DARPin E01, E68, E69	Designed Ankyrin Repeat protein; synthetic phage N3C DARPin library with signal recognition particle export selected on EGFR domain I-III fused to IgG1-Fc, E01 and E68 bind domain II, E69 binds domain I		[67]
Adnectin 1	Fibronectin tenth type III domain protein; synthetic library selected on EGFR-ECD-Fc by mRNA display; binds domain I		[84]
E6.2.6; EI3.4.3, A', D'	Fibronectin type III domain protein; synthetic library selected on EGFR 404SG produced in yeast by yeast display and cell sorting; ^{64}Cu-DOTA-fibronectin used for PET imaging		[82,130]
rAC1 Repebody	Variable lymphocyte receptor of jawless vertebrates derived scaffold; synthetic library selected on EGFR-ECD by phage display; binds domain III; chemo-enzymatically coupled to monomethyl auristatin F		[72]
Bi-specific formats			
MDX-447 EMD 82633	Bispecific/crosslinked of F(ab') from mAb H22 binding CD64/FcγRI and F(ab') from humanized mAb 425	I; D	[131]
EGFR/CD2 C_HC_L Miniantibody	scFv 425 and anti CD2 scFv M1 heterodimerized by fusion to IgG C_H1 and C_L, respectively		[51]
EGFR/CD2 DiBi Miniantibody	scFv 425 and anti CD2 scFv M1 fused to the N- and C-terminus of a helical dimerization domain resulting in a dimeric bispecific construct		[50]
BsALFH, BsFLAH (11F8, 2F8)	IgG-like tetravalent bispecific mAbs composed of anti EGFR scFv 11F8 and anti IGFR1 scFv 2F8 fused to C_L or C_H1-C_H2-C_H3, respectively		[132]
Di-Diabody IgG (11F8, A12)	Anti-EGFR Fv IMC-11F8 and anti-IGFR IMC-A12 Fv diabody fused to IgG Fc		[133]
C-BiTE and P-BiTE	Bispecific T-cell engager, anti-CD33 scFv fused to scFv derived from Cetuximab or scFv derived from Panitumumab, respectively		[52]
bscEGFRvIIIxCD3	Bispecific T-cell engager, anti-CD33 scFv fused to anti-EGFRvIII scFv MR1-1		[134]
DL11f (Fab-	Two-in-one antibody with Fv specificity for EGFR and Erb3; selection		[54]

ACCEPTED MANUSCRIPT

Format) MEHD7945A (Ig-Format)	of a heavy chain only CDR Fab library for blocking TGFα binding to EGFR yielded D1.5 which was CDR light chain randomized and selected for ErbB3 binding by phage display		
E01_GS_E69; E01_LZ1_E69	Bispecific DARPin binding two epitopes on EGFR made by fusion of E01 and E69 either by a GlySer-linker or a homodimerizing leucine zipper/coiled coil domain		[68]
CONAN-1 (7D12 and 9G8)	COoperative NANobody comprising 7D12 and 9G8 binding two epitopes on EGFR and albumin binding nanobody Alb1 fused via a 30 aa flexible GlySer linker; nanobodies obtained by phage selection of a library by competitive elution on EFGR-ECD; binding competes for mAb425 or Cetuximab, respectively		[92]
E#1-GS10-I-PEG and E#4-GS10-I-PEG	Tandem/bispecific Adnectin comprising anti-EGFR (E) and anti-IGFR (I) Adnectin linked by a GlySer-linker; Adnectins obtained by mRNA selection		[83]
MetHer1	2+1 antibody comprinsing cetuximab Fv grafted on both heteromeric knobs-into-holes Fc IgG1 and one anti-MET scFv from mAb 5D5 (onartuzumab) fused to one Fc part		[135]
ME22S	mAb-scFv antibody comprising Anti-Met Ab SAIT301 fused C-terminally with a disulfide stabilized anti-EGFR scFv obtained from a synthetic scFv library by phage display		[136]
FL518	four-in-one antibody combining anti-EGFR and anti-erbB3 two-in-one antibody Fab MEHD7945A via a knobs-into-holes Fc with two-in-one anti-HER2 anti-VEGF Fab bH1-44		[137]
CRTB6	tetra-specific antibody combining dual Fv domain, CrossMAb and knobs-into-holes technology; the Fv from anti-Her2 Trastuzumab on top of the Fv from Cetuximab Fv is combined via an knobs-into-holes Fc with the anti-erbB3 Fv from RG7116 and anti-VEGF Fv from Bevacizumab		[137]
CT9, CB12, RT6	mAbs based on CrossMab and knobs- into-holes technology combining the specificities of Trastuzumab×Cetuximab, Cetuximab×Bevacizumab and Trastuzumab×RG7116 respectively		[137]
EGFRBi	Cetuximab chemical herteroconjugated with anti-CD33 mAb OKT3	I	[138,139]
Bival1, Bipar1, Bipar2	Bivalent and bispecific/biparatopic VHHs linked by a 10aa GlySer linker; Bival1 (7D12-7D12), and Bipar1 (7D12–9G8), Bipar2 (9G8–EGb4) binding two eptiopes on EGFR		[93]
Ligand based (recent examples)			
EGF-Dianthin	Ribosome inhibiting toxin genetically fused to EGF		[140]
EGF-PAMAM-DNA	Charge based self-assembling nano particles composed of EGF, DNA, and polyamidoamine		[141]
EGF-FTH1	EGF genetically fused to human ferritin H-chain protein, which forms nano particels		[142]
Peptides			
GE11	12 aa peptide obtained by selecting a commercial phage display peptide library on EGFR		[96]
D4	6 aa peptide obtained by computational design on a EGFR surface pocket, used for liposome targeting		[98]
2a-4c; EDA1-EDA6	7-8 aa peptides; EGFR dimerization arm mimetics stabilized by a selenylsulfide or a triazolyl bridge		[100,143]

ACCEPTED MANUSCRIPT

Inherbin3	17 aa peptide derived from the dimerization arm of ErbB3 (244-LVYNKLTFQLEPNPHTK-260) also blocks EGFR:EGF stimulation		[144]
Clone 1 - 4	7 aa peptides; obtained from a phage display library selected on EGFR		[97]
Peptide 1 - 28	Variations of cysteine bridged dimerization arm mimetics most with 8 aa wt hairpin structure		[99, 145]
EC	The C-loop of epidermal growth factor (22 amino acids of EGF COOH terminal) genetically fused to toxins		[146]
PEPCH1 EGFRVIII	15 aa peptide with complement hydopathy to the EGFRvIII deletion junction		[147]
CVRAC; $_{D}$(CVRAV)	5 aa peptide binding EGF ligand and its retro inverso peptide that binds EGFR		[148]
Cp28	Peptidomimetic low affinity binder of the ligand EGF		[149]
ERRP	EGFR-related protein from rat can block EGFR activity when expressed		[150]
Disruptin	6 aa peptide from the cytosolic domain, which blocks dimerization and promotes degradation in transfection experiements		[101]
Viral capsids and adaptors			
DARPIn-VP2 rAAV; Affibody-VP2 rAAV	DARPin_E01 and Affibody_$Z_{EGFR:1907}$ genetically fused to VP2 of Adeno-associated virus with natural tropism knock-down for a virus directed enzyme prodrug approach		[151]
oncolytic measles viruses	scFv 425 genetically fused to mutated H in a full length viral genome		[152]
Adenobody	Anti-Adenovirus fiber-protein scFv genetically fused to EGF for retargeting Adenovirus		[153]
sCAR-EGF	EGF genetically fused to soluble coxsackie-Ad receptor (CAR) for Adenovirus retargeting		[154]
DARPin scFV Adapter	Trimerized bispecific DARPin comprising an anti-Adenovirus-knob protein, a trimerization motif with linker and DARPin_E01 or DARPin_E69		[155]
ESV1	EGFR-specific virus based on Sindbis virus (SINV) displaying EGF fused to E2, optimized by tandem selection and enrichment system (TSES)		[156]
Ad5.GE11.KO1	Adenovirus with peptide GE11 inserted in HI loop and natural tropism mutation		[157]
MR1-1-gCΔ HSV	Herpes simplex virus (HSV) produced with scFv MR1-1 genetically fused to gC deleted for natural binding		[158]
MHVsoR-EGF	coronavirus murine hepatitis virus (MHV) produced with genetically encoded EGF fused to mCEACAM, a soluble receptor		[110]
AdFK4m:Ecoil-EGF	Adenovirus with tropism mutations carrying K-coil peptide at fiber C-terminus interacts with E-coil peptide fused to EGF		[159]
Ad5-E1/AFP-E3/NIS:PAMAM-G2-PEG-GE11	Adenovirus was coated with cationic poly(amidoamine) (PAMAM) dendrimer linked to peptide GE11 for delivery of iodine enriching channel gene		[160]
Cet-LV	Lentivirus with streptavidin fused to vesicular stomatitis virus G protein were conjugated to biotinylated EGF or ^{111}Indium labeled mAb Cetuximab		[161,162]
HSV1716EGFR	Herpes simplex virus (HSV) 1716 variant displaying a scFv derived from mAB108		[108]

ACCEPTED MANUSCRIPT

	Via fusion to 274–393 of gD		

[a] Clinic phase: A: approved in major markets; (A): approved in smaller markets or limited approval; I, II, III: testing in clinical phase I to III; D: further clinical development discontinued as hinted on web pages

References

Reference annotations:

**of interest*

***of considerable interest*

1. Lemmon MA, Schlessinger J. Cell signaling by receptor tyrosine kinases. Cell. 2010;141(7), 1117–1134.

2. Roskoski R. The ErbB/HER family of protein-tyrosine kinases and cancer. Pharmacol. Res. 2014;79, 34–74.

3. Roskoski R. ErbB/HER protein-tyrosine kinases: Structures and small molecule inhibitors. Pharmacol. Res. 2014;87, 42–59.

4. Ullrich A, Coussens L, Hayflick J, *et al.* Human epidermal growth factor receptor cDNA sequence and aberrant expression of the amplified gene in A431 epidermoid carcinoma cells. Nature. 1984;309(5967), 418–425.

5. Ogiso H, Ishitani R, Nureki O, *et al.* Crystal structure of the complex of human epidermal growth factor and receptor extracellular domains. Cell. 2002;110(6), 775–787.

6. Arkhipov A, Shan Y, Das R, *et al.* Architecture and Membrane Interactions of the EGF Receptor. Cell. 2013;152(3), 557–569. **Extensive computational simulation studies of EFGR dimerization and activation with special attention to the membrane and transmembrane

ACCEPTED MANUSCRIPT

helices highlights intricate protein membrane interaction.

7. Hynes NE, Lane H a. ERBB receptors and cancer: the complexity of targeted inhibitors. Nat. Rev. Cancer. 2005;5(5), 341–54.

8. Bessman NJ, Bagchi A, Ferguson KM, Lemmon MA. Complex Relationship between Ligand Binding and Dimerization in the Epidermal Growth Factor Receptor. Cell Rep. 2014;9(4), 1306–1317.

9. Burgess AW, Cho H, Eigenbrot C, *et al.* An open-and-shut case? Recent insights into the activation of EGF/ErbB receptors. Mol. Cell. 2003;12(3), 541–552.

10. Ferguson KM, Berger MB, Mendrola JM, *et al.* EGF Activates Its Receptor by Removing Interactions that Autoinhibit Ectodomain Dimerization. Mol. Cell. 2003;11, 507–517.

11. Chung I, Akita R, Vandlen R, Toomre D, Schlessinger J, Mellman I. Spatial control of EGF receptor activation by reversible dimerization on living cells. Nature. 2010;464(7289), 783–787.

12. Arkhipov A, Shan Y, Das R, *et al.* Architecture and Membrane Interactions of the EGF Receptor. Cell. 2013;152(3), 557–569.

13. Liu P, Cleveland TE, Bouyain S, Byrne PO, Longo P a., Leahy DJ. A single ligand is sufficient to activate EGFR dimers. Proc. Natl. Acad. Sci. 2012;109(27), 10861–10866. **

The crystal structure of ErbB4/HER4 extracellular region complexed with Neuregulin-1β also shows asymmetric dimers similar to that Drosophila EGFR:spitz and in combination with further cellular data suggest that activity is attained in the single ligated state.

14. Schlessinger J. Ligand-induced, receptor-mediated dimerization and activation of EGF receptor. Cell. 2002;110(6), 669–672.

15. Jura N, Zhang X, Endres NF, Seeliger MA, Schindler T, Kuriyan J. Catalytic Control in the EGF

ACCEPTED MANUSCRIPT

Receptor and Its Connection to General Kinase Regulatory Mechanisms. Mol. Cell. 2011;42(1), 9–22.

16. Jura N, Shan Y, Cao X, Shaw DE, Kuriyan J. Structural analysis of the catalytically inactive kinase domain of the human EGF receptor 3. Proc. Natl. Acad. Sci. U. S. A. 2009;106(51), 21608–21613.

17. Wilson KJ, Gilmore JL, Foley J, Lemmon MA, Riese DJ. Functional selectivity of EGF family peptide growth factors: Implications for cancer. Pharmacol. Ther. 2009;122(1), 1–8.

18. Endres NF, Das R, Smith AW, *et al.* Conformational coupling across the plasma membrane in activation of the EGF receptor. Cell. 2013;152(3), 543–556. *

Several synthetic EGFR derived constructs were studied by fluorescence microscopy and in combination with transmembrane analyses the findings corroborate that EGF binding removes constraints and promotes activation through N-terminal association of the transmembrane helices.

19. Jura N, Endres NF, Engel K, *et al.* Mechanism for Activation of the EGF Receptor Catalytic Domain by the Juxtamembrane Segment. Cell. 2009;137(7), 1293–1307.

20. Doerner A, Scheck R, Schepartz A. Growth Factor Identity Is Encoded by Discrete Coiled-Coil Rotamers in the EGFR Juxtamembrane Region. Chem. Biol. 2015;22(6), 776–784.

21. Yarden Y, Sliwkowski MX. Untangling the ErbB signalling network. Nat. Rev. Mol. Cell Biol. 2001;2(2), 127–137.

22. Engelman JA. Targeting PI3K signalling in cancer: opportunities, challenges and limitations. Nat. Rev. Cancer. 2009;9(8), 550–562.

23. Vanhaesebroeck B, Stephens L, Hawkins P. PI3K signalling: the path to discovery and understanding. Nat. Rev. Mol. Cell Biol. 2012;13(3), 195–203.

ACCEPTED MANUSCRIPT

24. Kadamur G, Ross EM. Mammalian phospholipase C. Annu Rev Physiol. 2013;75, 127–154.

25. Alwan HAJ, van Zoelen EJJ, van Leeuwen JEM. Ligand-induced lysosomal epidermal growth factor receptor (EGFR) degradation is preceded by proteasome-dependent EGFR de-ubiquitination. J. Biol. Chem. 2003;278(37), 35781–35790.

26. Polo S. Signaling-mediated control of ubiquitin ligases in endocytosis. BMC Biol. 2012;10(25), 1–9.

27. Harris RC, Chung E, Coffey RJ. EGF receptor ligands. Exp. Cell Res. 2003;284, 2–13.

28. Blobel CP. ADAMs: key components in EGFR signalling and development. Nat. Rev. Mol. Cell Biol. 2005;6(1), 32–43.

29. Fischer OM, Hart S, Gschwind A, Ullrich A. EGFR signal transactivation in cancer cells. Biochem. Soc. Trans. 2003;, 1203–1208.

30. Friedländer E, Barok M, Szöllosi J, Vereb G. ErbB-directed immunotherapy: antibodies in current practice and promising new agents. Immunol. Lett. 2008;116(2), 126–140.

31. Nicholson RI, Gee JMW, Harper ME. EGFR and cancer prognosis. Eur. J. Cancer. 2001;37, S9–S15.

32. Hutchinson RA, Adams RA, McArt DG, Salto-Tellez M, Jasani B, Hamilton PW. Epidermal growth factor receptor immunohistochemistry: new opportunities in metastatic colorectal cancer. J. Transl. Med. 2015;13(217), 1–11.

33. FDA EMA. http://www.cancer.gov/about-cancer/treatment/drugs/cetuximab http://www.cancer.gov/about-cancer/treatment/drugs/panitumumab http://www.cancer.gov/about-cancer/treatment/drugs/necitumumab http://www.ema.europa.eu/ema/. 2016;

34. Masui H, Kawamoto T, Sato JD, Wolf B, Sato G, Mendelsohn J. Growth inhibition of human

ACCEPTED MANUSCRIPT

tumor cells in athymic mice by anti- epidermal growth factor receptor monoclonal antibodies. Cancer Res. 1984;44(3), 1002–1007.

35. Li S, Schmitz KR, Jeffrey PD, Wiltzius JJW, Kussie P, Ferguson KM. Structural basis for inhibition of the epidermal growth factor receptor by cetuximab. Cancer Cell. 2005;7(4), 301–311.

36. Rocha-Lima CM, Soares HP, Raez LE, Singal R. Special Report EGFR Targeting of Solid Tumors. Cancer Control. 1998;2, 295–304.

37. Schmiedel J, Blaukat A, Li S, *et al.* Matuzumab binding to EGFR prevents the conformational rearrangement required for dimerization. Cancer Cell. 2008;13(4), 365–373.

38. Yang X, Jia X, Corvalan JRF, Wang P, Davis CG, Jakobovits A. Eradication of Established Tumors by a Fully Human Monoclonal Antibody to the Epidermal Growth Factor Receptor without Concomitant hemotherapy Eradication of Established Tumors by a Fully Human Monoclonal Antibody to the Epidermal Growth Factor Receptor w. Cancer Res. 1999;15(59), 1236–1243.

39. Haigler H, Ash JF, Singer SJ, Cohen S. Visualization by fluorescence of the binding and internalization of epidermal growth factor in human carcinoma cells A-431. Proc. Natl. Acad. Sci. U. S. A. 1978;75(7), 3317–3321.

40. Yang XD, Jia XC, Corvalan JRF, Wang P, Davis CG. Development of ABX-EGF, a fully human anti-EGF receptor monoclonal antibody, for cancer therapy. Crit. Rev. Oncol. Hematol. 2001;38(1), 17–23.

41. Li S, Kussie P, Ferguson KM. Structural basis for EGF receptor inhibition by the therapeutic antibody IMC-11F8. Structure. 2008;16(2), 216–227.

42. Kenanova V, Olafsen T, Crow DM, *et al.* Tailoring the pharmacokinetics and positron emission tomography imaging properties of anti-carcinoembryonic antigen single-chain Fv-Fc antibody fragments. Cancer Res. 2005;65(2), 622–631.

ACCEPTED MANUSCRIPT

43. Shah DK. Pharmacokinetic and pharmacodynamic considerations for the next generation protein therapeutics. J. Pharmacokinet. Pharmacodyn. 2015;42(5), 553–571.

44. Pack P, Müller K, Zahn R, Plückthun A. Tetravalent miniantibodies with high avidity assembling in Escherichia coli. J. Mol. Biol. 1995;246(1), 28–34.

45. Holliger P, Hudson PJ. Engineered antibody fragments and the rise of single domains. Nat. Biotechnol. 2005;23(9), 1126–1136.

46. Nelson AL. Antibody fragments: Hope and hype. MAbs. 2010;2(1), 77–83.

47. Omidfar K, Shirvani Z. Single domain antibodies: a new concept for epidermal growth factor receptor and EGFRvIII targeting. DNA Cell Biol. 2012;31(6), 1015–1026.

48. Farajnia S, Ahmadzadeh V, Tanomand A, Veisi K, Khosroshahi SA, Rahbarnia L. Development trends for generation of single-chain antibody fragments. Immunopharmacol. Immunotoxicol. 2014;36(5), 297–308.

49. Trejtnar F, Laznicek M. Analysis of renal handling of radiopharmaceuticals. Q. J. Nucl. Med. 2002;46(3), 181–194.

50. Müller KM, Arndt KM, Plückthun A. A dimeric bispecific miniantibody combines two specificities with avidity. FEBS Lett. 1998;432(1-2), 45–49.

51. Müller KM, Arndt KM, Strittmatter W, Plückthun A. The first constant domain (C(H)1 and C(L)) of an antibody used as heterodimerization domain for bispecific miniantibodies. FEBS Lett. 1998;422(2), 259–264.

52. Lutterbuese R, Raum T, Kischel R, *et al.* T cell-engaging BiTE antibodies specific for EGFR potently eliminate KRAS- and BRAF-mutated colorectal cancer cells. Proc. Natl. Acad. Sci. U. S. A. 2010;107(28), 12605–12610. *

Cetuximab- and panitumumab derived scFvs were combined with an anti-CD3 scFv (BiTE) and shown

ACCEPTED MANUSCRIPT

to engage T- cells albeit in a narrow therapeutic window in a colorectal cancer xenograft models.

53. Bostrom J, Yu S-F, Kan D, *et al.* Variants of the antibody herceptin that interact with HER2 and VEGF at the antigen binding site. Science. 2009;323(5921), 1610–1614.

54. Schaefer G, Haber L, Crocker LM, *et al.* A Two-in-One Antibody against HER3 and EGFR Has Superior Inhibitory Activity Compared with Monospecific Antibodies. Cancer Cell. 2011;20(4), 472–486. *

A two-in-one bitargeting single Fv domain against EGFR and ErbB3 was created by phage display, reformatted as IgG, and shown to have superior therapeutic properties compared to that of the anti-ErbB3 antibody.

55. Brezski RJ, Georgiou G. Immunoglobulin isotype knowledge and application to Fc engineering. Curr. Opin. Immunol. 2016;40(May 2016), 62–69.

56. Liu G, Tu D, Lewis M, *et al.* Fc-γ Receptor Polymorphisms, Cetuximab Therapy, and Survival in the NCIC CTG CO.17 Trial of Colorectal Cancer. Clin. Cancer Res. 2016;22(10), 2435–2444.

57. Akash MSH, Rehman K, Parveen A, Ibrahim M. Antibody-drug conjugates as drug carrier systems for bioactive agents. Int. J. Polym. Mater. Polym. Biomater. 2016;65(1), 1–10.

58. Chari RVJ, Miller ML, Widdison WC. Antibody-drug conjugates: An emerging concept in cancer therapy. Angew. Chemie - Int. Ed. 2014;53(15), 3796–3827.

59. Kim EG, Kim KM. Strategies and advancement in antibody-drug conjugate optimization for targeted cancer therapeutics. Biomol. Ther. 2015;23(6), 493–509.

60. Koefoed K, Steinaa L, Søderberg JN, *et al.* Rational identification of an optimal antibody mixture for targeting the epidermal growth factor receptor. MAbs. 2011;3(6), 584–95.

61. Kearns JD, Bukhalid R, Sevecka M, *et al.* Enhanced Targeting of the EGFR Network with MM-

ACCEPTED MANUSCRIPT

151, an Oligoclonal Anti-EGFR Antibody Therapeutic. Mol. Cancer Ther. 2015;14(7), 1625–1636.

62. Arena S, Siravegna G, Mussolin B, *et al.* MM-151 overcomes acquired resistance to cetuximab and panitumumab in colorectal cancers harboring EGFR extracellular domain mutations. Sci. Transl. Med. 2016;8(324ra14), 1–10.

63. Main ERG, Lowe AR, Mochrie SGJ, Jackson SE, Regan L. A recurring theme in protein engineering: the design, stability and folding of repeat proteins. Curr. Opin. Struct. Biol. 2005;15(4), 464–471.

64. Plückthun A. Designed Ankyrin Repeat Proteins (DARPins): Binding Proteins for Research, Diagnostics, and Therapy. Annu. Rev. Pharmacol. Toxicol. 2015;55, 489–511.

65. Forrer P, Stumpp MT, Binz HK, Plückthun A. A novel strategy to design binding molecules harnessing the modular nature of repeat proteins. FEBS Lett. 2003;539(1-3), 2–6.

66. Interlandi G, Wetzel SK, Settanni G, Plückthun A, Caflisch A. Characterization and Further Stabilization of Designed Ankyrin Repeat Proteins by Combining Molecular Dynamics Simulations and Experiments. J. Mol. Biol. 2008;375(3), 837–854.

67. Steiner D, Forrer P, Plückthun A. Efficient selection of DARPins with sub-nanomolar affinities using SRP phage display. J. Mol. Biol. 2008;382(5), 1211–1227.

68. Boersma YL, Chao G, Steiner D, Wittrup KD, Plückthun A. Bispecific Designed Ankyrin Repeat Proteins (DARPins) targeting epidermal growth factor receptor inhibit A431 cell proliferation and receptor recycling. J. Biol. Chem. 2011;286(48), 41273–41285.

69. Boersma YL, Plückthun A. DARPins and other repeat protein scaffolds: Advances in engineering and applications. Curr. Opin. Biotechnol. 2011;22(6), 849–857.

70. Enkhbayar P, Kamiya M, Osaki M, Matsumoto T, Matsushima N. Structural Principles of

ACCEPTED MANUSCRIPT

Leucine-Rich Repeat (LRR) Proteins. Proteins Struct. Funct. Genet. 2004;54(3), 394–403.

71. Lee S-CS-G, Park K, Han J, *et al.* Design of a binding scaffold based on variable lymphocyte receptors of jawless vertebrates by module engineering. Proc. Natl. Acad. Sci. 2012;109(9), 3299–3304.

72. Lee JJ, Choi HJ, Yun M, *et al.* Enzymatic Prenylation and Oxime Ligation for the Synthesis of Stable and Homogeneous Protein-Drug Conjugates for Targeted Therapy. Angew. Chemie - Int. Ed. 2015;54(41), 12020–12024. *

A repebody based on the repeat protein internalin B was selected against EGFR, co-crystallized and conjugated to a drug by a chemo-enzymatic approach.

73. Sihver W, Pietzsch J, Krause M, Baumann M, Steinbach J, Pietzsch HJ. Radiolabeled cetuximab conjugates for EGFR targeted cancer diagnostics and therapy. Pharmaceuticals. 2014;7(3), 311–338.

74. Cuartero-Plaza A, Mart??nez-Miralles E, Rosell R, Vadell-Nadal C, Farr?? M, Real FX. Radiolocalization of squamous lung carcinoma with 131I-labeled epidermal growth factor. Clin. Cancer Res. 1996;2(1), 13–20.

75. Friedman M, Nordberg E, Hoiden-Guthenberg I, *et al.* Phage display selection of Affibody molecules with specific binding to the extracellular domain of the epidermal growth factor receptor. Protein Eng. Des. Sel. 2007;20(4), 189–199.

76. Löfblom J, Frejd FY, Ståhl S. Non-immunoglobulin based protein scaffolds. Curr. Opin. Biotechnol. 2011;22(6), 843–848.

77. Justino CIL, Duarte AC, Rocha-Santos TAP. Analytical applications of affibodies. TrAC - Trends Anal. Chem. 2015;65, 73–82.

78. Tolmachev V, Rosik D, Wållberg H, *et al.* Imaging of EGFR expression in murine xenografts

ACCEPTED MANUSCRIPT

using site-specifically labelled anti-EGFR 111In-DOTA-Z EGFR:2377 Affibody molecule: aspect of the injected tracer amount. Eur. J. Nucl. Med. Mol. Imaging. 2010;37(3), 613–622.

79. Dickinson CD, Veerapandian B, Dai XP, *et al.* Crystal structure of the tenth type III cell adhesion module of human fibronectin. J. Mol. Biol. 1994;236(4), 1079–1092.

80. Main AL, Harvey TS, Baron M, Boyd J, Campbell ID. The three-dimensional structure of the tenth type III module of fibronectin: An insight into RGD-mediated interactions. Cell. 1992;71(4), 671–678.

81. Koide A, Bailey CW, Huang X, Koide S. The fibronectin type III domain as a scaffold for novel binding proteins. J. Mol. Biol. 1998;284(4), 1141–1151.

82. Hackel BJ, Ackerman ME, Howland SW, Wittrup KD. Stability and CDR Composition Biases Enrich Binder Functionality Landscapes. J. Mol. Biol. 2010;401(1), 84–96.

83. Emanuel SL, Engle LJ, Chao G, *et al.* A fibronectin scaffold approach to bispecific inhibitors of epidermal growth factor receptor and insulin-like growth factor-I receptor. MAbs. 2011;3(1), 38–48.

84. Ramamurthy V, Krystek SR, Bush A, *et al.* Structures of adnectin/protein complexes reveal an expanded binding footprint. Structure. 2012;20(2), 259–269. *

An fibronectin monobody was selected against EGFR, the crystral structure of the complex was determined, and the binding probed by alanine scanning.

85. Hackel BJ, Kimura RH, Gambhir SS. Use of 64Cu-labeled Fibronectin Domain with EGFR-Overexpressing Tumor Xenograft: Molecular Imaging. Radiology. 2012;263(1), 179–188.

86. Yan Y, Chen G, Wei H, *et al.* Fast Photochemical Oxidation of Proteins (FPOP) Maps the Epitope of EGFR BInding to Adnectin. J. Am. Soc. Mass Spectrom. 2014;25(12), 2084–2092. *

The FPOP technique was used to identify a binding surface.

ACCEPTED MANUSCRIPT

87. Revets H, De Baetselier P, Muyldermans S. Nanobodies as novel agents for cancer therapy. Expert Opin. Biol. Ther. 2005;5(1), 111–124.

88. Arbabi Ghahroudi M, Desmyter A, Wyns L, Hamers R, Muyldermans S. Selection and identification of single domain antibody fragments from camel heavy-chain antibodies. FEBS Lett. 1997;414(3), 521–526.

89. Kijanka M, Dorresteijn B, Oliveira S, van Bergen en Henegouwen PMP. Nanobody-based cancer therapy of solid tumors. Nanomedicine (Lond). 2015;10(1), 161–174.

90. Roovers RC, Laeremans T, Huang L, *et al.* Efficient inhibition of EGFR signaling and of tumour growth by antagonistic anti-EFGR Nanobodies. Cancer Immunol. Immunother. 2007;56(3), 303–317.

91. Schmitz KR, Bagchi A, Roovers RC, van Bergen en Henegouwen PMP, Ferguson KM. Structural Evaluation of EGFR Inhibition Mechanisms for Nanobodies/VHH Domains. Structure. 2013;21(7), 1214–1224. *

Co-crystal structures of EGFR with the VHHs 7D12, EgA1, and 9G8 illuminate interactions and the modes of inhibition.

92. Roovers RC, Vosjan MJWD, Laeremans T, *et al.* A biparatopic anti-EGFR nanobody efficiently inhibits solid tumour growth. Int. J. Cancer. 2011;129(8), 2013–2024.

93. Heukers R, Vermeulen JF, Fereidouni F, *et al.* Endocytosis of EGFR requires its kinase activity and N-terminal transmembrane dimerization motif. J. Cell Sci. 2013;126(21), 4900–4912.

94. Van Der Meel R, Oliveira S, Altintas I, *et al.* Tumor-targeted Nanobullets: Anti-EGFR nanobody-liposomes loaded with anti-IGF-1R kinase inhibitor for cancer treatment. J. Control. Release. 2012;159(2), 281–289.

95. Oliveira S, Schiffelers RM, van der Veeken J, *et al.* Downregulation of EGFR by a novel

ACCEPTED MANUSCRIPT

multivalent nanobody-liposome platform. J. Control. Release. 2010;145(2), 165–175.

96. Li Z, Zhao R, Wu X, *et al.* Identification and characterization of a novel peptide ligand of epidermal growth factor receptor for targeted delivery of therapeutics. FASEB J. 2005;19(14), 1978–1985.

97. Nakamura T, Takasugi H, Aizawa T, *et al.* Peptide mimics of epidermal growth factor (EGF) with antagonistic activity. J. Biotechnol. 2005;116(3), 211–219.

98. Song S, Liu D, Peng J, *et al.* Novel peptide ligand directs liposomes toward EGF-R high-expressing cancer cells in vitro and in vivo. FASEB J. 2009;23, 1396–1404.

99. Mizuguchi T, Ohara N, Iida M, *et al.* Evaluation of dimerization-inhibitory activities of cyclic peptides containing a β-hairpin loop sequence of the EGF receptor. Bioorganic Med. Chem. 2012;20(19), 5730–5737. *

Various peptides including retro-inverso versions were derived from the dimerzation arm and evaluated for their EGFR-dimerization inhibiting properties.

100. Hanold LE, Oruganty K, Ton NT, Beedle AM, Kannan N, Kennedy EJ. Inhibiting EGFR Dimerization Using Triazolyl-Bridged Dimerization Arm Mimics. PLoS One. 2015;10(3), e0118796.

101. Ahsan A, Ray D, Ramanand SG, *et al.* Destabilization of the epidermal growth factor receptor (EGFR) by a peptide that inhibits EGFR binding to heat shock protein 90 and receptor dimerization. J. Biol. Chem. 2013;288(37), 26879–26886.

102. Collet G, Grillon C, Nadim M, Kieda C. Trojan horse at cellular level for tumor gene therapies. Gene. 2013;525(2), 208–216.

103. Jain PK, El-Sayed IH, El-Sayed MA. Au nanoparticles target cancer. Nano Today. 2007;2(1), 18–29.

ACCEPTED MANUSCRIPT

104. Altintas I, Heukers R, Van Der Meel R, *et al.* Nanobody-albumin nanoparticles (NANAPs) for the delivery of a multikinase inhibitor 17864 to EGFR overexpressing tumor cells. J. Control. Release. 2013;165(2), 110–118.

105. Li W, Liu Z, Li C, *et al.* Radionuclide therapy using 131I-labeled anti-epidermal growth factor receptor-targeted nanoparticles suppresses cancer cell growth caused by EGFR overexpression. J. Cancer Res. Clin. Oncol. 2016;142(3), 619–632.

106. Mamot C, Drummond DC, Greiser U, *et al.* Epidermal Growth Factor Receptor (EGFR)-targeted Immunoliposomes Mediate Specific and Efficient Drug Delivery to EGFR- and EGFRvIII-overexpressing Tumor Cells 1. Control. 2003;(6), 3154 –3161.

107. Hagen S, Baumann T, Wagner HJ, *et al.* Modular adeno-associated virus (rAAV) vectors used for cellular virus-directed enzyme prodrug therapy. Sci. Rep. 2014;4, 3759.

108. Conner J, Braidwood L, Brown SM. A strategy for systemic delivery of the oncolytic herpes virus HSV1716: redirected tropism by antibody-binding sites incorporated on the virion surface as a glycoprotein D fusion protein. Gene Ther. 2008;15(24), 1579–1592.

109. Piao Y, Jiang H, Alemany R, *et al.* Oncolytic adenovirus retargeted to Delta-EGFR induces selective antiglioma activity. Cancer Gene Ther. 2009;16(3), 256–65.

110. Verheije MH, Lamfers MLM, Würdinger T, *et al.* Coronavirus genetically redirected to the epidermal growth factor receptor exhibits effective antitumor activity against a malignant glioblastoma. J. Virol. 2009;83(15), 7507–7516.

111. Goldstein NI, Prewett M, Zuklys K, Rockwell P, Mendelsohn J. Biological efficacy of a chimeric antibody to the epidermal growth factor receptor in a human tumor xenograft model. Clin. Cancer Res. 1995;1(11), 1311–1318.

112. Mateo C, Moreno E, Amour K, Lombardero J, Harris W, P??rez R. Humanization of a mouse monoclonal antibody that blocks the epidermal growth factor receptor: Recovery of

ACCEPTED MANUSCRIPT

antagonistic activity. Immunotechnology. 1997;3(1), 71–81.

113. Bleeker WK, van Bueren JJ, van Ojik HH, *et al.* Dual mode of action of a human anti-epidermal growth factor receptor monoclonal antibody for cancer therapy. J. Immunol. 2004;173(7), 4699–4707.

114. Murthy U, Basu A, Rodeck U, Herlyn M, Ross AH, Das M. Binding of an antagonistic monoclonal antibody to an intact and fragmented EGF-receptor polypeptide. Arch. Biochem. Biophys. 1987;252(2), 549–560.

115. Wersäll P, Ohlsson I, Biberfeld P, *et al.* Intratumoral infusion of the monoclonal antibody, mAb 425, against the epidermal-growth-factor receptor in patients with advanced malignant glioma. Cancer Immunol. Immunother. 1997;44(3), 157–164.

116. Kettleborough C a, Saldanha J, Heath VJ, Morrison CJ, Bendig MM. Humanization of a mouse monoclonal antibody by CDR-grafting: the importance of framework residues on loop conformation. Protein Eng. 1991;4(7), 773–783.

117. Dean C, Modjtahedi H, Eccles S, Box G, Styles J. Immunotherapy with antibodies to the EGF receptor. Int. J. Cancer. Suppl. 1994;8, 103–107.

118. Gerdes CA, Nicolini VG, Herter S, *et al.* GA201 (RG7160): A novel, humanized, glycoengineered anti - EGFR antibody with enhanced ADCC and superior in vivo efficacy compared with cetuximab. Clin. Cancer Res. 2013;19(5), 1126–1138.

119. Panousis C, Rayzman VM, Johns TG, *et al.* Engineering and characterisation of chimeric monoclonal antibody 806 (ch806) for targeted immunotherapy of tumours expressing de2-7 EGFR or amplified EGFR. Br. J. Cancer. 2005;92(6), 1069–1077.

120. Phillips AC, Boghaert ER, Vaidya KS, *et al.* ABT-414, an Antibody-Drug Conjugate Targeting a Tumor-Selective EGFR Epitope. Mol. Cancer Ther. 2016;128(2), 438–445.

ACCEPTED MANUSCRIPT

121. Zalutsky MR, Boskovitz A, Kuan CT, *et al.* Radioimmunotargeting of malignant glioma by monoclonal antibody D2C7 reactive against both wild-type and variant III mutant epidermal growth factor receptors. Nucl. Med. Biol. 2012;39(1), 23–34.

122. Chandramohan V, Bao X, Keir ST, *et al.* Construction of an immunotoxin, D2C7-(scdsFv)-PE38KDEL, targeting EGFRwt and EGFRvIII for brain tumor therapy. Clin. Cancer Res. 2013;19(17), 4717–4727.

123. Winkler ME, O'Connor L, Winget M, Fendly B. Epidermal growth factor and transforming growth factor alpha bind differently to the epidermal growth factor receptor. Biochemistry. 1989;28(15), 6373–6378.

124. Chao G, Cochran JR, Dane Wittrup K. Fine epitope mapping of anti-epidermal growth factor receptor antibodies through random mutagenesis and yeast surface display. J. Mol. Biol. 2004;342(2), 539–550.

125. Gooi HC, Hounsell EF, Lax I, *et al.* The carbohydrate specificities of the monoclonal antibodies 29.1, 455 and 3C1B12 to the epidermal growth factor receptor of A431 cells. Biosci. Rep. 1985;5(1), 83–94.

126. Lorimer IA, Keppler-Hafkemeyer A, Beers RA, Pegram CN, Bigner DD, Pastan I. Recombinant immunotoxins specific for a mutant epidermal growth factor receptor: targeting with a single chain antibody variable domain isolated by phage display. Proc. Natl. Acad. Sci. U. S. A. 1996;93(25), 14815–14820.

127. Huang L, Gainkam LOT, Caveliers V, *et al.* SPECT imaging with 99mTc-labeled EGFR-specific nanobody for in vivo monitoring of EGFR expression. Mol. Imaging Biol. 2008;10(3), 167–175.

128. Nevoltris D, Lombard B, Dupuis E, Mathis G, Chames P, Baty D. Conformational nanobodies reveal tethered epidermal growth factor receptor involved in EGFR/ErbB2 predimers. ACS Nano. 2015;9(2), 1388–1399.

ACCEPTED MANUSCRIPT

129. Friedman M, Orlova A, Johansson E, *et al.* Directed evolution to low nanomolar affinity of a tumor-targeting epidermal growth factor receptor-binding affibody molecule. J. Mol. Biol. 2008;376(5), 1388–13402.

130. Hackel BJ, Sathirachinda A, Gambhir SS. Designed hydrophilic and charge mutations of the fibronectin domain: Towards tailored protein biodistribution. Protein Eng. Des. Sel. 2012;25(10), 639–647.

131. Fury MG, Lipton A, Smith KM, Winston CB, Pfister DG. A phase-I trial of the epidermal growth factor receptor directed bispecific antibody MDX-447 without and with recombinant human granulocyte-colony stimulating factor in patients with advanced solid tumors. Cancer Immunol. Immunother. 2008;57(2), 155–163.

132. Lu D, Zhang H, Ludwig D, *et al.* Simultaneous Blockade of Both the Epidermal Growth Factor Receptor and the Insulin-like Growth Factor Receptor Signaling Pathways in Cancer Cells with a Fully Human Recombinant Bispecific Antibody. J. Biol. Chem. 2004;279(4), 2856–65.

133. Lu D, Zhang H, Koo H, *et al.* A fully human recombinant IgG-like bispecific antibody to both the epidermal growth factor receptor and the insulin-like growth factor receptor for enhanced antitumor activity. J. Biol. Chem. 2005;280(20), 19665–19672.

134. Choi BD, Kuan C-T, Cai M, *et al.* Systemic administration of a bispecific antibody targeting EGFRvIII successfully treats intracerebral glioma. Proc. Natl. Acad. Sci. U. S. A. 2013;110(1), 270–275.

135. Castoldi R, Ecker V, Wiehle L, *et al.* A novel bispecific EGFR/Met antibody blocks tumor-promoting phenotypic effects induced by resistance to EGFR inhibition and has potent antitumor activity. Oncogene. 2013;32(50), 5593–5601.

136. Lee JM, Lee SH, Hwang J-W, *et al.* Novel strategy for a bispecific antibody: induction of dual target internalization and degradation. Oncogene. 2016;advance on, 1–10.

ACCEPTED MANUSCRIPT

137. Hu S, Fu W, Xu W, *et al.* Four-in-one antibodies have superior cancer inhibitory activity against EGFR, HER2, HER3, and VEGF through disruption of HER/MET crosstalk. Cancer Res. 2015;75(1), 159–170. **

Several multispecific whole antibody formats were generated and evaluated for their pharmaceutical potential.

138. Reusch U, Sundaram M, Davol PA, *et al.* Anti-CD3 × anti-epidermal growth factor receptor (EGFR) bispecific antibody redirects T-cell cytolytic activity to EGFR-positive cancers in vitro and in an animal model. Clin. Cancer Res. 2006;12(1), 183–190.

139. Zitron IM, Thakur A, Norkina O, Barger GR, Lum LG, Mittal S. Targeting and killing of glioblastoma with activated T cells armed with bispecific antibodies. BMC Cancer. 2013;13(83), 1–14.

140. Mallinckrodt B von, Thakur M, Weng A, *et al.* Dianthin-EGF is an effective tumor targeted toxin in combination with saponins in a xenograft model for colon carcinoma. Futur. Oncol. 2014;10(14), 2161–2175.

141. Li J, Chen L, Liu N, Li S, Hao Y, Zhang X. EGF-coated nano-dendriplexes for tumor-targeted nucleic acid delivery in vivo. Drug Deliv. 2015;early onli, 1–8.

142. Li X, Qiu L, Zhu P, *et al.* Epidermal growth factor-ferritin H-chain protein nanoparticles for tumor active targeting. Small. 2012;8(16), 2505–2514.

143. Hanold LE, Watkins CP, Ton NT, Liaw P, Beedle AM, Kennedy EJ. Design of a selenylsulfide-bridged EGFR dimerization arm mimic. Bioorganic Med. Chem. 2015;23(12), 2761–2766.

144. Xu R, Povlsen GK, Soroka V, Bock E, Berezin V. A peptide antagonist of the ErbB1 receptor inhibits receptor activation, tumor cell growth and migration in vitro and xenograft tumor growth in vivo. Cell. Oncol. 2010;32(4), 259–274.

ACCEPTED MANUSCRIPT

145. Mizuguchi T, Uchimura H, Kakizawa T, *et al.* Inhibitory effect of a dimerization-arm-mimetic peptide on EGF receptor activation. Bioorganic Med. Chem. Lett. 2009;19(12), 3279–3282.

146. Liu WJ, Liu XJ, Li L, Li Y, Zhang SH, Zhen YS. Tuftsin-based, EGFR-targeting fusion protein and its enediyne-energized analog show high antitumor efficacy associated with CD47 down-regulation. Cancer Immunol. Immunother. 2014;63(12), 1261–1272.

147. Campa MJ, Kuan CT, O'Connor-McCourt MD, Bigner DD, Patz EF. Design of a novel small peptide targeted against a tumor-specific receptor. Biochem. Biophys. Res. Commun. 2000;275(2), 631–636.

148. Cardó-Vila M, Giordano RJ, Sidman RL, *et al.* From combinatorial peptide selection to drug prototype (II): targeting the epidermal growth factor receptor pathway. Proc. Natl. Acad. Sci. U. S. A. 2010;107(11), 5118–5123.

149. Guardiola S, Díaz-Lobo M, Seco J, García J, Nevola L, Giralt E. Peptides targeting EGF block the EGF-EGFR interaction. Chembiochem. 2015;10.1002/cb.

150. Yu Y, Rishi a K, Turner JR, *et al.* Cloning of a novel EGFR-related peptide: a putative negative regulator of EGFR. Am. J. Physiol. Cell Physiol. 2001;280(5), C1083–C1089.

151. Hagen S, Baumann T, Wagner HJ, *et al.* Modular adeno-associated virus (rAAV) vectors used for cellular virus-directed enzyme prodrug therapy. Sci. Rep. 2014;4(3759), 1–11.

152. Nakamura T, Peng K-W, Harvey M, *et al.* Rescue and propagation of fully retargeted oncolytic measles viruses. Nat. Biotechnol. 2005;23(2), 209–214.

153. Watkins SJ, Mesyanzhinov V V., Kurochkina LP, Hawkins RE. The "adenobody" approach to viral targeting: specific and enhanced adenoviral gene delivery. Gene Ther. 1997;4(10), 1004–1012.

154. Hemminki A, Dmitriev I, Liu B, Desmond RA, Alemany R, Curiel DT. Targeting oncolytic

ACCEPTED MANUSCRIPT

adenoviral agents to the epidermal growth factor pathway with a secretory fusion molecule. Cancer Res. 2001;61(17), 6377–6381.

155. Dreier B, Honegger A, Hess C, *et al.* Development of a generic adenovirus delivery system based on structure-guided design of bispecific trimeric DARPin adapters. Proc. Natl. Acad. Sci. U. S. A. 2013;110(10), E869–E877. *

A bispecifc trimeric DARPin construct was used to retarget a viral particle.

156. Dai H-S, Liu Z, Jiang W, Kuhn RJ. Directed evolution of a virus exclusively utilizing human epidermal growth factor receptor as the entry receptor. J. Virol. 2013;87(20), 11231–43.

157. Uusi-Kerttula H, Legut M, Davies J, *et al.* Incorporation of Peptides Targeting EGFR and FGFR1 into the Adenoviral Fiber Knob Domain and Their Evaluation as Targeted Cancer Therapies. Hum. Gene Ther. 2015;26(5), 320–329.

158. Grandi P, Fernandez J, Szentirmai O, *et al.* Targeting HSV-1 virions for specific binding to epidermal growth factor receptor-vIII-bearing tumor cells. Cancer Gene Ther. 2010;17(9), 655–663.

159. Zeng Y, Pinard M, Jaime J, *et al.* A ligand-pseudoreceptor system based onde novo designed peptides for the generation of adenoviral vectors with altered tropism. J. Gene Med. 2008;10(4), 355–367.

160. Grünwald GK, Vetter A, Klutz K, *et al.* EGFR-Targeted Adenovirus Dendrimer Coating for Improved Systemic Delivery of the Theranostic NIS Gene. Mol. Ther. Nucleic Acids. 2013;2(April), e131.

161. Huhtala T, Kaikkonen MU, Lesch HP, Viitala S, Ylä-Herttuala S, Närvänen A. Biodistribution and antitumor effect of Cetuximab-targeted lentivirus. Nucl. Med. Biol. 2014;41(1), 77–83.

162. Kaikkonen MU, Lesch HP, Pikkarainen J, *et al.* (Strept)avidin-displaying lentiviruses as versatile tools for targeting and dual imaging of gene delivery. Gene Ther. 2009;16(7), 894–904.

ACCEPTED MANUSCRIPT

Title:

AAV production in suspension: evaluation of different cell culture media and scale-up potential

Authors:

Rebecca C. Feiner, Kathrin Teschner, Irina Schierbaum, Julian Teschner, Kristian M. Müller

This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. An extraction of *BMC Proceedings***,** 12, 3 (2018) is reprinted in this work.

P-349

AAV production in suspension: evaluation of different cell culture media and scale-up potential

Rebecca C. Feiner, Kathrin Teschner, Irina Schierbaum, Julian Teschner, Kristian M. Müller
Cellular and Molecular Biotechnology, Bielefeld University, 33602 Bielefeld, Germany
Correspondence: Kristian M. Müller (kristian@syntbio.net)
BMC Proceedings 2018, **12(Suppl 1):**P-349

Background

Recombinant adeno-associated virus (rAAV) approaches have an outstanding reputation in gene therapy and are evaluated for cancer therapy [1]. Advantages include long-term gene expression, targeting of dividing and non-dividing cells, and low immunogenicity. Established rAAV production utilizes triple transfection of adherent HEK 293 cells, which hardly meets product yield requirements for clinical applications. We transferred the AAV production system to HEK 293-F suspension cells. This process is scalable and uses serum-free media streamlining downstream procedures. After optimization of transfection efficiencies and shaker cultivations, we produced titers of 1×10^5 viral genomes per cell in a 2 l bioreactor.

Materials and methods

The suspension adapted HEK-FreeStyle 293-F cell line was used for the experiments in chemically defined animal component free media (HEK-TF, HEK-GM (Xell AG), Freestyle F17 (Thermo Fisher Scientific)). Samples for viable cell density and viabilities were taken daily and analyzed using an automated cell counting system (CEDEX, Roche Diagnostics). Transient transfection of 3×10^6 cells/ml was carried out with polyethylenimine Max in a 1:4 DNA-PEI ratio (w/w) with 2 µg DNA. Three plasmids (pGOI, pRepCap, pHelper) were applied in a molar 1:1:1 ratio (Fig. 1a). Pretests were performed in orbital shaking tube spin bioreactors. For scale-up, batch processes were carried out in 125 ml shake flasks as well as in 2 l stirred bioreactors at 30% air saturation and pH 7.1. Transfection efficiencies and rAAV production were quantified by flow cytometry using a GOI coding for a fluorescent protein and qPCR of genomic copies, respectively.

Results

By optimizing the DNA amount for transfection of 293-F cells more than 90 % of the cells were reproducibly transfected. Batch cultivations in shaker flasks revealed that rAAV were produced in the first 24-96 h after transfection. Figure 1b shows viable cell densities and viabilities in relation to the genomic titer. Genomic titers were determined from raw cell extracts and up to 10^9 copies/ml were repetitively achievable. A decrease in viability marked the decline in genomic copies per ml showing that a prolongation of the process e.g. by addition of a feed would probably not increase yield. In a first scale-up, the rAAV production was transferred to a 2 l bioreactor (Fig. 1c). Transfection efficiencies in bioreactors of up to 55% were comparable to that obtained in a simultaneous shaker flask experiment. Transfection efficiencies were lower compared to prior experiments due to controlled conditions in the bioreactor. Nonetheless the titer with up to 1×10^5 genomic copies per cell was elevated compared to that of shaker flasks.

Conclusions

First experiments with 293-F cells in HEK TF medium showed promising results of transferring rAAV production from the adherent system to suspension. After improvement of transfections by the adjustment of DNA amounts in small scale experiments, AAV production was analyzed in shaker flasks. The batch process showed an expected increase in cell density with low variability between biological replicates (Fig. 1b). The genomic titer increased according to the viable cell density until day four where a sudden drop started. This observation was made for AAV productions in HEK-TF, HEK-GM and Freestyle F17 medium. For optimal yields, we assume that a slight decrease in viability marks the point in time for harvest. From optimized protocols, a batch process in a 2 l bioreactor was carried out. Interestingly the bioreactor cultivation resulted in lower overall viable cell densities but in higher genomic copies per cell compared to shaker flasks (Fig. 1c). These results are comparable to already published data for suspension cells [2]. Subsequent optimization of the bioreactor protocol will lead to further increase in rAAV yield.

Acknowledgements

The authors thank Xell AG, Bielefeld, for providing HEK serum-free media (HEK GM and HEK TF) and for fruitful discussions.

References

1. Hagen S, Baumann T, Wagner HJ, Morath V, Kaufmann B, Fischer A, et al. **Modular adeno-associated virus (rAAV) vectors used for cellular virus-directed enzyme prodrug therapy.** *Sci Rep*. 2014;**4**:3759.
2. Grieger JC, Soltys SM, Samulski RJ. **Production of Recombinant Adeno-associated Virus Vectors Using Suspension HEK293 Cells and Continuous Harvest of Vector From the Culture Media for GMP FIX and FLT1 Clinical Vector.** *Mol Ther*. 2016;**24**:287–97.

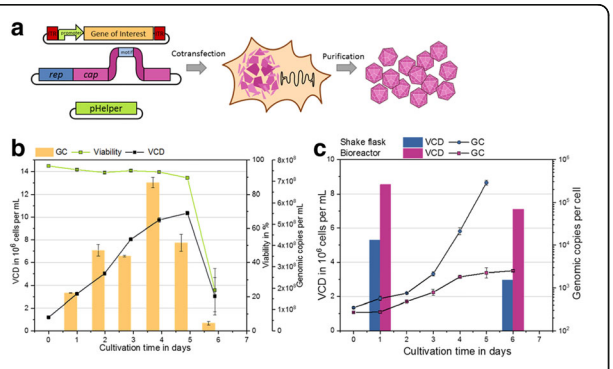


Fig. 1 (abstract P-349). a Schematic overview of rAAV production in HEK293 cells with triple-transfection system. **b** Viable cell densities (VCD), viabilities and genomic copies per ml (GC) of a rAAV production with 293-F batch cultivations in shaker flasks. Genomic copies per ml refer to the titer determined in 1 ml culture volume. Error bars represent biological and technical duplicate measurements of samples. **c** Viable cell densities and genomic copies per cell of a rAAV production with 293-F batch cultivation in a 2 l bioreactor. For reasons of comparability between shaker and bioreactor data genomic copies are given per cell. Error bars represent technical duplicate measurements of samples

9 Appendix: Publications

Title:

EGF-mCherry Fusion Protein Expressed in E. coli Shows Product Heterogeneity but a High Biological Activity

Authors:

Rebecca C. Feiner, Ina Pennè, Benjamin Müller, and Kristian M. Müller

Reprinted with permission from *Biochemistry* 2019, 58, 8, 1043-1047. Copyright 2019 American Chemical Society.

Biochemistry

Communication

Cite This: Biochemistry 2019, 58, 1043−1047

pubs.acs.org/biochemistry

EGF-mCherry Fusion Protein Expressed in *E. coli* Shows Product Heterogeneity but a High Biological Activity

Rebecca C. Feiner,[†] Ina Pennè,[†] Benjamin Müller,[‡] and Kristian M. Müller*[,†]

[†]Cellular and Molecular Biotechnology, Faculty of Technology, Bielefeld University, Bielefeld 33613, Germany

[‡]Biofidus AG, Hainteichstrasse 78, Bielefeld 33613, Germany

Supporting Information

ABSTRACT: The epidermal growth factor receptor (EGFR) is a transmembrane protein involved in cell signaling processes, and dysregulation of its activity often drives tumor growth. EGFR is a clinically validated tumor marker and target for antibodies and tyrosine kinase inhibitors. We demonstrate that a fusion protein of the natural ligand epidermal growth factor (EGF) with the fluorescent reporter mCherry can be expressed in the cytosol of *E. coli* in high yields and with a high biological activity. Biophysical characterization by mass spectrometry analysis confirmed three disulfide bonds that are crucial for protein structure. Biolayer interferometry data of the protein−protein interaction of EGF-mCherry with the soluble EGFR are comparable to that of unmodified EGF. Cell culture experiments demonstrated that this fusion replicates all important features of the natural ligand. Finally, fluorescent assays based on EGF-mCherry provided a simple and convenient method to compare EGFR levels on cells and to determine competition of EGFR-binding molecules. These assays will help to rank competitive properties of EGFR inhibitors.

Epidermal growth factor (EGF) is one of seven main natural ligands of the epidermal growth factor receptor (EGFR), also named ErbB1 as member of the ErbB receptor family.[1,2] EGF is synthesized as a type I transmembrane protein with an extracellular N-terminus and a cytosolic C-terminus.[3] In the transmembrane form, EGF binds juxtaposed EGFR. Enzymatic cleavage with disintegrin and metalloproteases releases mature EGF.[4] EGF is a 53 amino acid polypeptide of about 6 kDa.[5] Six cysteines within the chain form three internal disulfide bonds, which are important for biological activity.[6] The ErbB family is ubiquitously expressed in epithelial, mesenchymal and neuronal cells and their cellular progenitors.[7] According to the standard model, EGFR dimerizes upon ligand binding inducing intracellular tyrosine kinase activity and phosphorylation. Cellular responses vary with the type of ligand and the type of receptor dimer pair.[8] EGFR signaling is involved in cell growth and migration as well as proliferation and differentiation and the receptor was shown to be overexpressed in different types of cancer.[9] To analyze expression levels of EGFR, internalization, and inhibition of ligands, mostly [^{125}I]-labeled EGF is used.[10] An advantage of this system is the high detection sensitivity, which contrasts with the handling demands and short half-life of the radioactivity. Using a fluorescent protein coupled to EGF is environmentally friendly and compatible with modern fluorescent workflows. The 28.8 kDa monomer mCherry is a modification of the fluorescent protein DsRed from the organism *Discosoma sp*, which maintains fluorescence under acidic conditions.[11] This offers the opportunity to trace mCherry during internalization. EGFR internalization is clathrin-dependent, and the forming vesicles fuse with early endosomes.[12] A pH stable reporter benefits tracing of the receptor until the final degradation in lysosomes. A variant of EGF with a fluorescent reporter fused to its N-terminus and expressed in *E. coli* (eGFP-EGF) was already described by Jiang et al.,[13] who also analyzed colocalization with the EGFR. However, detailed aspects of the fusion protein have not yet been discussed. We close this gap with further data on biological activity and biophysical characterization.

According to long-standing publications, EGF expressed in the cytosol of *E. coli* is mainly sequestered to inclusion bodies.[14] In the reducing environment of the *E. coli*, cytosol proteins are not expected to form disulfide bonds, if not enforced by structural proximity. Protein fusions seem to increase solubility, and recent publications showed that EGF fusions can be expressed in the cytosol of *E. coli*.[15] As these reports lack a characterization of the disulfide state and a comparison with a third party human EGF (hEGF), we expand on this work and additionally test a fusion to the C-terminus of EGF with a more acid-stable fluorescent protein.

We constructed a fusion of EGF with mCherry attached to the C-terminus at the genetic level in a standard pET21a vector. Crystal structure analyses and the cell-bound state of full EGF indicate that the C-terminus of mature EGF is equally if not better suited for insertions than previous N-terminal fusions as it is more distant to a disulfide bridge (Figures S1 and S2). The fusion protein was expressed in *E. coli* BL21(DE3) and purified via immobilized metal ion affinity chromatography (IMAC) (Figure 1A) followed by anion exchange chromatography (Figure S3A). Good yields of EGF-mCherry have been obtained (14.9 mg per 500 mL shaker culture in LB). Only a small amount of EGF-mCherry was found in inclusion bodies in SDS-PAGE analysis. We assume that the well folding, larger mCherry, which comes off the ribosome second, prevents the EGF from improper folding, facilitates the adoption of the correct 3D structure, and enables

Received: January 9, 2019
Revised: February 3, 2019
Published: February 8, 2019

ACS Publications © 2019 American Chemical Society

DOI: 10.1021/acs.biochem.9b00021
Biochemistry 2019, 58, 1043−1047

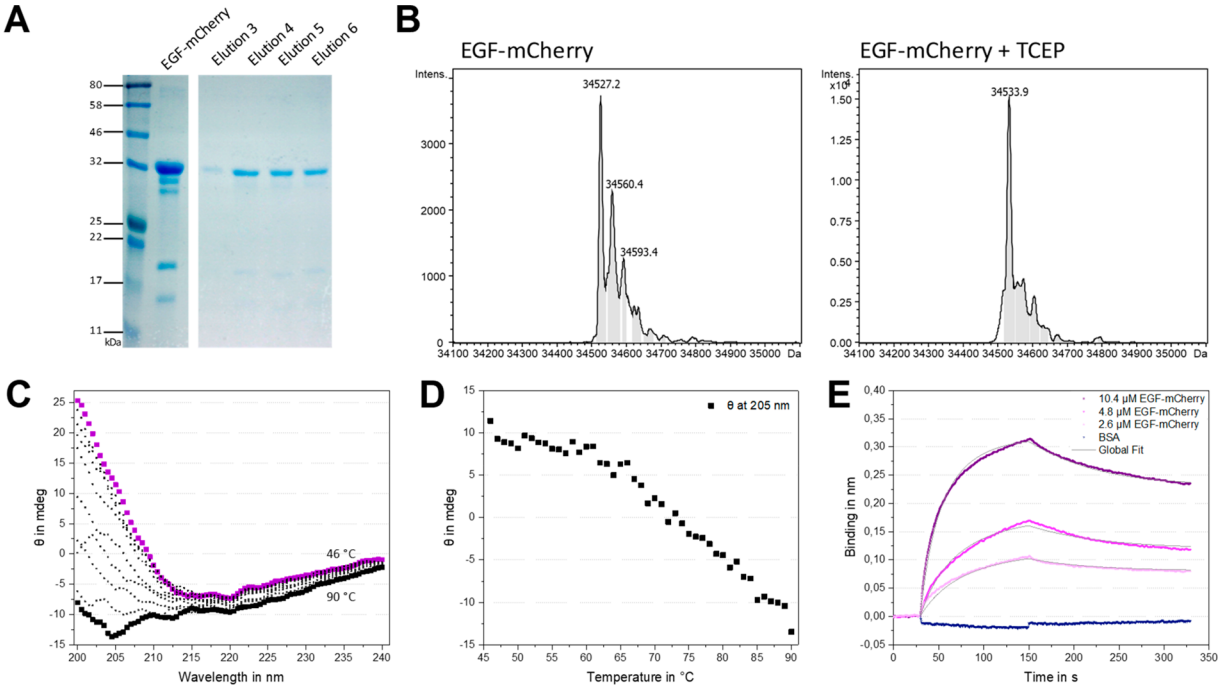


Figure 1. Characterization of EGF-mCherry. (A) EGF-mCherry was purified using Ni-NTA affinity chromatography (second lane) followed by an ion exchange chromatography (lanes 3−6). (B) LC-ESI-MS analysis of EGF-mCherry under nonreducing and reducing conditions. (C, D) Circular dichroism spectra of a temperature course experiment from 46 to 90 °C in 1 °C steps. (E) A soluble EGFR was immobilized on an ARG2 sensor for biolayer interferometry (BLI). Different concentrations of EGF-mCherry were measured, and the data was fitted.

purification from the soluble fraction. Disulfide bonds may form during production or upon cell disruption and purification. In the final product, disulfide bonds were analyzed using mass spectrometry. The reduced and nonreduced EGF-mCherry samples were analyzed in a LC-ESI-TOF mass spectrometer. Each disulfide bond results in a mass difference of 2 Da in the molecular weights of the reduced vs unreduced proteins. As a control, commercially available hEGF (Gibco) with tested biological activity was used. A shift of 6 Da is expected as three disulfide bridges should be reduced to thiols. This shift (6.7 Da) is approximately observed for EGF-mCherry (Figure 1B).

The two side peaks with +33 Da each are best explained by trisulfide formation at one or two disulfide spots or alternatively by methionine oxidation to a Sulphone.[16] For the purchased hEGF, only a 5 Da shift was observed (Figure S4). EGF-mCherry was further analyzed on a SDS gel under reducing and nonreducing conditions with and without boiling (Figure S5). For the non-DDT-nonboiled sample, partial formation of oligomers was observed, which were reduced upon DTT addition. Analysis of MS and SDS-PAGE data also revealed fragmentation at the chromophore N-acylimine at elevated temperatures.[17] Structural properties and thermal stability of EGF-mCherry were also determined with circular dichroism spectroscopy. Figure 1C−D presents measurements from 46 to 90 °C, demonstrating a high stability. Even at 90 °C, the spectrum shows remaining structural elements, which is in accordance with literature values for mCherry.[18]

The affinity of EGF to the soluble extracellular domain of EGFR (sEGFR residues 1−621) was investigated with biolayer interferometry (BLI). sEGFR was secreted from 293F suspension culture after transient transfection and purified via IMAC (Figure S3B). The receptor was immobilized on an amine-reactive biosensor (ARG2, BLItz system), and EGF-mCherry was measured at different concentrations. No binding was observed for the control bovine serum albumin (BSA), demonstrating specificity of the EGF-EGFR interaction. A dissociation constant K_D for EGF-mCherry of 7.6 nM was determined using a global fit based on a 1:1 kinetic binding model on three concentrations shown in Figure 1E. The value is in agreement with the 1.8 nM reported for surface plasmon resonance experiments of the EGF−EGFR interaction.[19] It is known that the 1:1 binding model does not adequately describe the interaction because different receptor forms are known to be present in solution.[20] Nevertheless, these data enable a comparison and show an interaction in the expected range corroborating that production of EGF in the cytosol of *E. coli* yields a biological active form.

In addition to biophysical properties, we were also interested in the biological activity. EGF is known to stimulate several cell signaling processes upon binding to the EGFR. These effects can be analyzed in wound healing assays. Here, the closure of a scratch mark in a monolayer of cells of the EGFR-overexpressing A431 cell line is evaluated. The size of the wound, which typically diminishes linearly, at different time intervals yields the cell migration rate (Figure 2).[21] A detailed analysis is presented in Figure S6. The cell migration rate for EGF-mCherry was 21.2 ± 1.6 μm/h. A commercial hEGF preparation showed a rate of 13.16 ± 0.22 μm/h at an equal molar concentration. Differences could be explained by multimeric EGF-mCherry forms that might contribute to a higher activity.

The EGF-mCherry fusion offers further applications in cell culture experiments. We expand previous work by showing that EGF-mCherry is a versatile tool to assess the relative EGFR surface concentration in a comparison of different cell lines. Cell lines expressing varying levels of EGFR were chosen to test whether fluorescence intensity increases with increasing

DOI: 10.1021/acs.biochem.9b00021
Biochemistry 2019, 58, 1043−1047

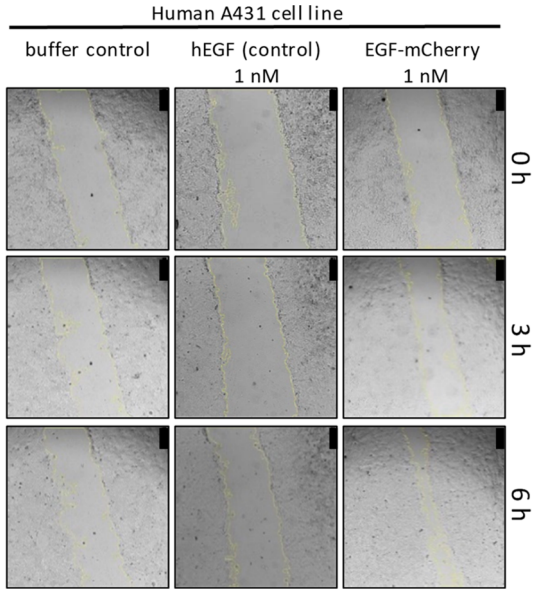


Figure 2. Wound healing assay using A431 cells. Cells incubated with 0 nM (buffer control) and 1 nM EGF-mCherry or a commercial hEGF. Cell-free areas and cell migration rates were determined using the ImageJ software plugin MRI Wound Healing Tool. Scale bars in the upper right corner represent 250 μm.

receptor density. The cancer cells lines A431, HeLa, and MCF7 present EGFR in decreasing concentrations. A431 were shown to present about 3.0×10^6 receptors per cells.[22] For HeLa and MCF7, cells values of 2.0×10^5 and 1.5×10^4, respectively, were reported.[23,24] As seen in Figure 3A, fluorescence microscopy after incubation with 5 nM EGF-mCherry visualized a strong EGF-mCherry signal for A431 cells, whereas the intensity of fluorescence decreased for HeLa and further for MCF7 cells. A more detailed analysis was performed using flow cytometry analysis (Figure 3B). Here a variety of cell lines expressing different EGFR levels were analyzed with an excess of EGF-mCherry. The data agrees with the values described in literature, and also allows for an assessment of EGF receptor density of unknown cell lines compared to those described in literature. A stably EGFR-expressing cell line generated in our lab (ZMB-HEK293-EGFR) presents a EGFR signal in the range of HeLa cells and about three times more than the parental HEK293. Figure 3C shows internalization in A431 cells incubated with 10 nM EGF-mCherry after 15 and 30 min, revealing internalization into vesicles as previously described for eGFP-EGF.[13]

EGF-mCherry was further tested in cellular competition assays, which are compatible with today's fluorescent high-throughput equipment. This will help to characterize binding inhibitors and blocking antibodies for the EGFR tumor target.[25,26] Incubation of A431 cells with EGF-mCherry at increasing concentrations at 37 °C for 15 min entails several processes such as receptor binding, internalization, and depletion. The measured signal thus reflects bound and internalized EGF-mCherry. In Figure 4A, an expected sigmoidal curve can be seen with signal or receptor saturation, respectively, at this time point. Fitting a sigmoidal function, the half-maximal effective concentration EC_{50} was determined to

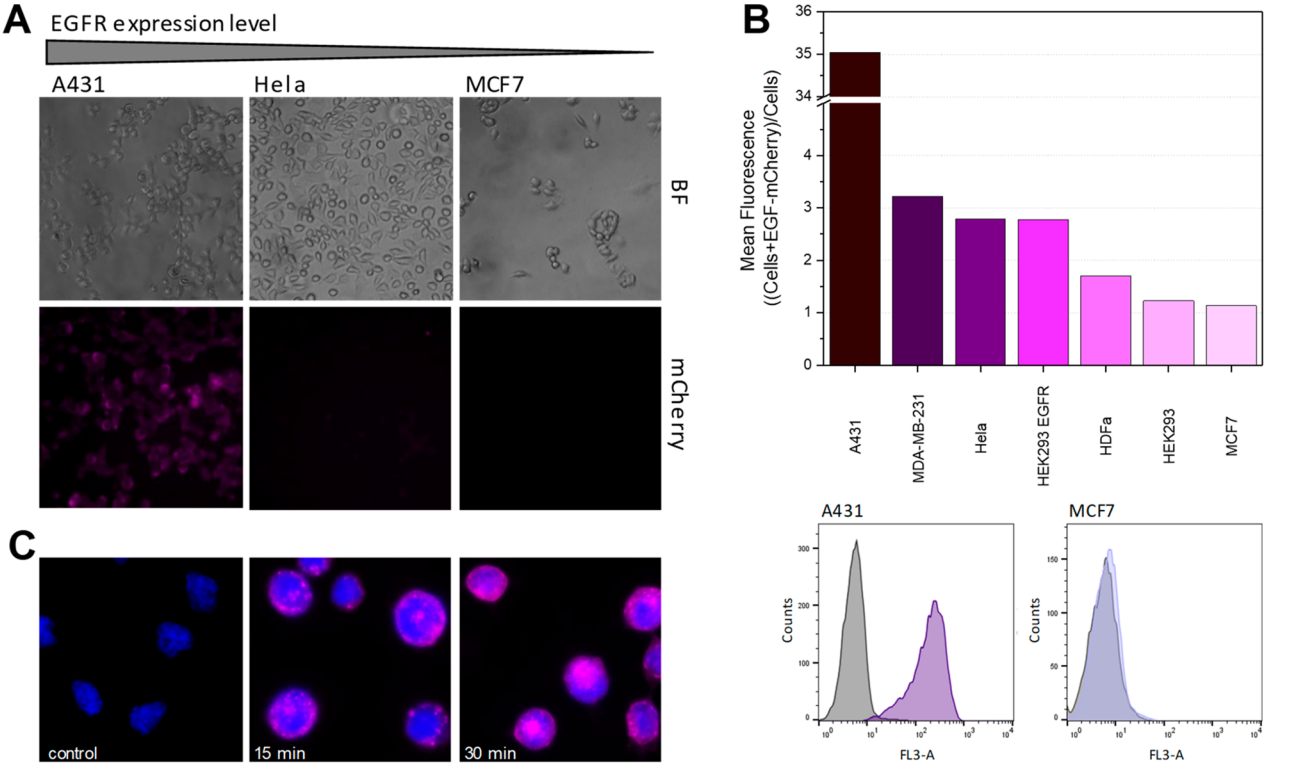


Figure 3. Characterization of EGF-mCherry in cell culture experiments. (A) A431, HeLa, and MCF7 cells were incubated with 5 nM EGF-mCherry for 15 min before fluorescence microscopy. (B) Flow cytometry analysis of cells expressing different levels of EGFR. Cells were incubated with 150 nM EGF-mCherry for 15 min before measurements. Folds of mean fluoresce difference with and without EGF-mCherry are presented for each cell line. (C) Internalization of EGF-mCherry was shown using A431 cells. Cells were incubated with 10 nM EGF-mCherry, fixed and counterstained with DAPI.

DOI: 10.1021/acs.biochem.9b00021
Biochemistry 2019, 58, 1043−1047

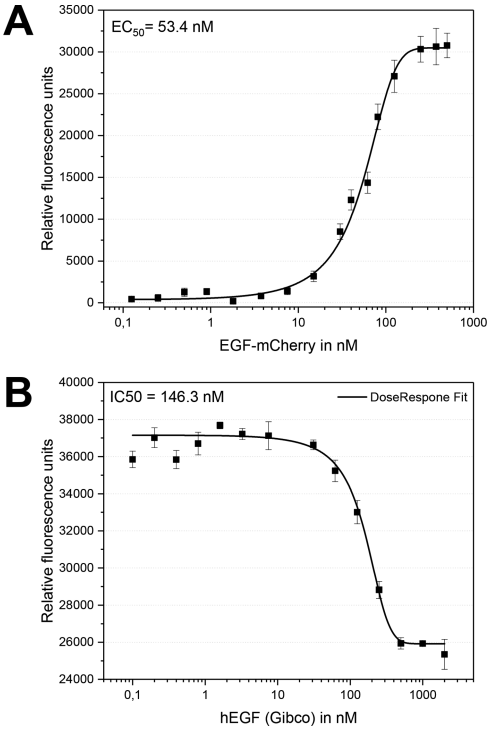


Figure 4. EGF-mCherry cell binding and competition assay. (A) A431 cells were incubated with an increasing concentration of EGF-mCherry. (B) Cells were simultaneously incubated with 50 nM EGF-mCherry and hEGF in increasing concentrations. Relative fluorescence of cells was measured using a Tecan Spark instrument at 550/610 nm.

be 53.4 ± 4.2 nM. This allowed estimation of the adequate EGF-mCherry concentration in subsequent competitive assays to be 50 nM, which was kept constant, while hEGF (Gibco) was added as a a competitor at increasing concentrations. A dose−response equation fit was used to determine an inhibitory concentration at half-maximum (IC_{50}) (Figure 4B). The IC_{50} value for hEGF was 146.3 ± 36.2 nM, which is almost three times the expected value, if both EGF would be equally effective. Differences in this value might be explained by the observed formation of multimers in EGF-mCherry and structural deficits in the hEGF. Competition assays based on this method are easy and simple to implement using standard laboratory equipment. Unlike commercial assays with radioactivity or special dyes, only a plate reader with an excitation in the range of 560 nm and emission above 580 nm is required.

In summary, we demonstrated cytosolic expression of EGF fused to the N-terminus of a fluorescent protein in *E. coli* and characterized its biophysical properties and product heterogeneity. Interaction of this EGF variant with soluble EGFR was demonstrated by BLI and with cell-bound EGFR by fluorescence microscopy and flow cytometry. Application in competition assays makes this protein interesting for analysis of any nonfluorescent protein interacting with the EGFR. Additionally, the fusion protein is also suitable for interaction measurements using FRET, for example, in combination with EGFP. Protein expression and purification are quick and simple using standard laboratory methods, and the results were consistent over different batches of EGF-mCherry productions (Figure S7).

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.biochem.9b00021.

Materials and methods, crystal structure analysis of EGF and EGF-mCherry, schematic overview of pET21a(+)-EGF-mCherry, SDS PAGE of EGF-mCherry and sEGFR after purification via IMAC, LC-ESI-MS analysis of EGF-mCherry and hEGF, SDS PAGE of EGF-mCherry under various conditions, wound healing assay, and EGF-mCherry competition assay (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: kristian@syntbio.net.

ORCID

Rebecca C. Feiner: 0000-0002-8784-0875

Kristian M. Müller: 0000-0002-7914-0625

Author Contributions

The manuscript was written with contributions of all authors. All authors have given approval to the final version of the manuscript.

Funding

This work was supported by grants from the Bielefeld University.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Philipp Borchert and Georg Falck for experimental assistance, Marco Radukic for providing the ZMB-HEK293-EGFR cell line, and Xell AG for providing HEK serum-free media (HEK-TF).

■ ABBREVIATIONS

EGF, epidermal growth factor; EGFR, epidermal growth factor receptor; TCEP, tris(2-carboxyethyl)phosphine; CD, circular dichroism; GFP, green fluorescent protein; BLI, biolayer interferometry

■ REFERENCES

(1) Cohen, S. (1962) Isolation of a Mouse Submaxillary Gland Protein Accelerating Incisor Eruption and Eyelid Opening in the New-Born Animal. *J. Biol. Chem. 237* (6), 1555−1562.

(2) Feiner, R. C., and Müller, K. M. (2016) Recent Progress in Protein-Protein Interaction Study for EGFR-Targeted Therapeutics. *Expert Rev. Proteomics 13* (9), 817−832.

(3) Massague, J., and Pandiella, A. (1993) Membrane-Anchored Growth Factors. *Annu. Rev. Biochem. 62* (1), 515−541.

(4) Fischer, O. M., Hart, S., Gschwind, A., and Ullrich, A. (2003) EGFR Signal Transactivation in Cancer Cells. *Biochem. Soc. Trans. 31* (6), 1203−1208.

(5) Savage, C. R., Inagami, T., and Cohen, S. (1972) The Primary Structure of Epidermal Growth Factor. *J. Biol. Chem. 247* (23), 7612−7621.

(6) Harris, R. C., Chung, E., and Coffey, R. J. (2003) EGF Receptor Ligands. *Exp. Cell Res. 284*, 2−13.

(7) Roskoski, R. (2014) The ErbB/HER Family of Protein-Tyrosine Kinases and Cancer. *Pharmacol. Res. 79*, 34−74.

(8) Yarden, Y., and Sliwkowski, M. X. (2001) Untangling the ErbB Signalling Network. *Nat. Rev. Mol. Cell Biol. 2* (2), 127−137.

DOI: 10.1021/acs.biochem.9b00021
Biochemistry 2019, 58, 1043−1047

(9) Nicholson, R. I., Gee, J. M. W., and Harper, M. E. (2001) EGFR and Cancer Prognosis. *Eur. J. Cancer 37*, S9−S15.
(10) Veale, D., Kerr, N., Gibson, G. J., and Harris, A. L. (1989) Characterization of Epidermal Growth Factor Receptor in Primary Human Non-Small Cell Lung Cancer. *Cancer Res. 49* (5), 1313−1317.
(11) Shaner, N. C., Campbell, R. E., Steinbach, P. A., Giepmans, B. N. G., Palmer, A. E., and Tsien, R. Y. (2004) Improved Monomeric Red, Orange and Yellow Fluorescent Proteins Derived from Discosoma Sp. Red Fluorescent Protein. *Nat. Biotechnol. 22* (12), 1567−1572.
(12) Alwan, H. A. J., van Zoelen, E. J. J., and van Leeuwen, J. E. M. (2003) Ligand-Induced Lysosomal Epidermal Growth Factor Receptor (EGFR) Degradation Is Preceded by Proteasome-Dependent EGFR de-Ubiquitination. *J. Biol. Chem. 278* (37), 35781−35790.
(13) Jiang, H., Zhang, J., Shi, B. Z., Xu, Y. H., Li, Z. H., and Gu, J. R. (2007) Application of EGFP-EGF Fusions to Explore Mechanism of Endocytosis of Epidermal Growth Factor. *Acta Pharmacol. Sin. 28* (1), 111−117.
(14) Allen, G., Paynter, C. A., and Winther, M. D. (1985) Production of Epidermal Growth Factor in Escherichia Coli from a Synthetic Gene. *J. Cell Sci. 3*, 29−38.
(15) Ma, Y., Yu, J., Lin, J., Wu, S., Li, S., and Wang, J. (2016) High Efficient Expression, Purification, and Functional Characterization of Native Human Epidermal Growth Factor in Escherichia Coli. *BioMed Res. Int. 2016*, 1.
(16) Nielsen, R. W., Tachibana, C., Hansen, N. E., and Winther, J. R. (2011) Trisulfides in Proteins. *Antioxid. Redox Signaling 15* (1), 67−75.
(17) Gross, L. A., Baird, G. S., Hoffman, R. C., Baldridge, K. K., and Tsien, R. Y. (2000) The Structure of the Chromophore within DsRed, a Red Fluorescent Protein from Coral. *Proc. Natl. Acad. Sci. U. S. A. 97* (22), 11990−11995.
(18) Borrell, K. L., Cancglin, C., Stinger, B. L., Defrates, K. G., Caputo, G. A., Wu, C., and Vaden, T. D. (2017) An Experimental and Molecular Dynamics Study of Red Fluorescent Protein MCherry in Novel Aqueous Amino Acid Ionic Liquids. *J. Phys. Chem. B 121* (18), 4823−4832.
(19) Kuo, W.-T., Lin, W.-C., Chang, K.-C., Huang, J.-Y., Yen, K.-C., Young, I.-C., Sun, Y.-J., and Lin, F.-H. (2015) Quantitative Analysis of Ligand-EGFR Interactions: A Platform for Screening Targeting Molecules. *PLoS One 10* (2), No. e0116610.
(20) Domagala, T., Konstantopoulos, N., Smyth, F., Jorissen, R. N., Fabri, L., Geleick, D., Lax, I., Schlessinger, J., Sawyer, W., Howlett, G. J., Burgess, A. W., and Nice, E. N. (2000) Stoichiometry, Kinetic and Binding Analysis of the Interaction between Epidermal Growth Factor (EGF) and the Extracellular Domain of the EGF Receptor. *Growth Factors 18* (1), 11−29.
(21) Jonkman, J. E. N., Cathcart, J. A., Xu, F., Bartolini, M. E., Amon, J. E., Stevens, K. M., and Colarusso, P. (2014) An Introduction to the Wound Healing Assay Using Live-Cell Microscopy. *Cell Adh. Migr. 8* (5), 440−451.
(22) Haigler, H., Ash, J. F., Singer, S. J., and Cohen, S. (1978) Visualization by Fluorescence of the Binding and Internalization of Epidermal Growth Factor in Human Carcinoma Cells A-431. *Proc. Natl. Acad. Sci. U. S. A. 75* (7), 3317−3321.
(23) Dickson, R. B., Hanover, J. A., Willingham, M. C., and Pastan, I. (1983) Prelysosomal Divergence of Transferrin and Epidermal Growth Factor during Receptor-Mediated Endocytosis. *Biochemistry 22* (24), 5667−5674.
(24) Reilly, R. M., Kiarash, R., Sandhu, J., Lee, Y. W., Cameron, R. G., Hendler, A., Vallis, K., and Gariépy, J. (2000) A Comparison of EGF and MAb 528 Labeled with 111In for Imaging Human Breast Cancer. *J. Nucl. Med. 41* (5), 903−911.
(25) Müller, K. M., Arndt, K. M., Strittmatter, W., and Plückthun, A. (1998) The First Constant Domain (CH1 and CL) of an Antibody Used as Heterodimerization Domain for Bispecific Miniantibodies. *FEBS Lett. 422* (2), 259−264.
(26) Müller, K. M., Arndt, K. M., and Plückthun, A. (1998) A Dimeric Bispecific Miniantibody Combines Two Specificities with Avidity. *FEBS Lett. 432* (1−2), 45−49.

DOI: 10.1021/acs.biochem.9b00021
Biochemistry 2019, 58, 1043−1047

EGF-mCherry Fusion Protein Expressed in *E. coli* Shows Product Heterogeneity but High Biological Activity

Rebecca C. Feiner,[†] Ina Pennè,[†] Benjamin Müller,[‡] Kristian M. Müller*[,†]

[†] Cellular and Molecular Biotechnology, Faculty of Technology, Bielefeld University, Bielefeld, Germany
[‡] Biofidus AG, Hainteichstrasse 78, 33613 Bielefeld, Germany
* Corresponding author: Kristian M. Müller, Cellular and Molecular Biotechnology, Faculty of Technology, Bielefeld University, 33615 Bielefeld, Germany; Email: kristian@syntbio.net

MATERIALS AND METHODS

Materials and Instrumentation. Reagents and chemicals were obtained at Sigma-Aldrich, VWR or Thermo Fisher Scientific. DNA oligonucleotides were ordered at Sigma-Aldrich. The BLI measurements were performed using a BLItz (ForteBio) instrument. DNA and Protein concentration were determined by absorbance at 260 and 280 nm respectively using a NanoDrop 2000c spectrophotometer.

Construction of pET21a-EGF-mCherry-His6. The human soluble EGF gene (UniProt ID P01133) was amplified by PCR from plasmid pZMB0008 with the primers (NheI_Met_EGF-for 5'-AAAAAGCTAG CATGAACAGC GACAGCGAGT GCCC-3' and GGSG_EGF_rev 5'-AGATCCTCCA CCAGATCCAC CACCCCTCAG CTCCCACCAC TTCAGG-3'). The mCherry gene (UniProt ID X5DSL3) was amplified by PCR from pZMB0267 with the primers GGSG_mCherry_for 5'-GGTGGTGGAT CTGGTGGAGG ATCTATGGTG TCCAAGGGCG AAGAGG-3' and mCherryHisStopXhoI_rev 5'-AAAACTCGAG TCAGTGGTGG TGGTGGTGGT GCTTGTACAG CTCATCCATG CCGC-3'. A third PCR reaction was used to fuse PCR product 1 and PCR product 2 to generate a DNA-fragment named NheI-Start-EGF-GGSG-mCherry-His6-Stop-XhoI. The NheI/XhoI digested fragment was integrated in the likewise opened pET21a vector to generate pZMB0490. Resulting inserts were analyzed for their correctness by Sanger DNA-sequencing (Sequencing Core Facility, CeBiTec, Bielefeld, Germany).

Construction of pcDNA6A EGFR-ECD-His6. The soluble EGF-receptor (UniProt ID P00533, residues 1-621) was provided in pcDNA6A for expression in 293F suspension culture. This plasmid was obtained by removing the EGFR transmembrane coding region within plasmid pcDNA6A-EGFR ECD (1-644), which was a gift from Mien-Chie Hung (Addgene plasmid #42666)[1], by amplifying the complete plasmid with primer FWD EGFR-TMD 5'-CTCGAGTCTA GAGGGCCC-3' and REV EGFR-TMD 5'-GGACGGGATC TTAGGCCC-3' with subsequent self-ligation. The insert in resulting plasmid pZMB0202 was analyzed for correctness by Sanger DNA sequencing (Sequencing Core Facility, CeBiTec, Bielefeld, Germany).

EGF-mCherry expression and purification. Plasmid pET21a-EGF-mCherry-His6 was transformed into competent *E. coli* BL21(DE3). An overnight culture (37 °C, 180 rpm) was used to inoculate a culture in 0.5 L LB medium containing 100 µg/mL ampicillin. This culture was grown (37 °C, 180 rpm) up to an OD_{600} of 0.5. Protein expression was then induced with isopropyl-β-D-thiogalactopyranoside (IPTG) (0.1 mM) and the cells were cultivated at 30 °C, 180 rpm for 4 h. Afterwards, cells were harvested by centrifugation (3220×*g*, 30 min), washed with 100 mM Na_2HPO_4 buffer (pH7.4) and stored at -20 °C. For protein purification, cells were thawed and resuspended in 30 mL equilibration buffer (50 mM Na_2HPO_4, 300 NaCl, 10 mM imidazole, pH 8). The cell membrane of bacteria was disrupted in three repeated French press cycles at a pressure of 6.9 mPa (1000 psi). Cell debris was separated from soluble protein by centrifugation at 15.000×g for 30 min at 4°C. The supernatant was filtered and applied to a 1 ml Protino Ni-NTA column for purification. After washing with 30 CV buffer containing 50 mM imidazole in equilibration buffer, the protein was eluted with an elution buffer containing 300 mM imidazole. Further

purification included an ion exchange chromatography (IEX) using an ÄKTA start protein purification system. The protein was applied to a self-packed Q Sepharose Fast Flow column (GE Healthcare) with a column volume of 1 ml and a gradient from 0 to 2000 mM NaCl in 50 mM Tris pH 8. Fractions with a volume of 1 ml were collected. Finally, eluted protein was concentrated to 14.9 mg/ml in 1× PBS using an Amicon Ultra centrifugal filter unit (MWCO 30 kDa).

Soluble EGFR expression and purification. The suspension adapted HEK-FreeStyle 293-F cell line (Thermo) was used for transient expression and cultivated in chemically defined HEK-TF (Xell AG) medium in 125 ml shake flasks with orbital shaking at 185 rpm and 5 cm amplitude at 37°C and 5% CO_2. Prior to transfection, 293F cells were washed with 1× PBS. Transient transfection of 3×10^6 cells/ml was carried out with polyethylenimine Max (PEI) (Polysciences) by mixing 2 µg DNA with 8 µg PEI (1:4 DNA-PEI ratio) in 250 µl HEK-TF for the transfection of 1 ml cells. 4 h post transfection 1 ml medium was added per 1.25 ml culture. Transfections were carried out in 125 ml shake flasks with initial 30 ml cell suspension. The cultivation of cells lasted until the viability of cells dropped below 60 %. Cells were harvested at 2000×g for 5 min, the cell culture supernatant was sterile filtered and applied to subsequent Ni-NTA purification as describe above.

CD spectroscopy. Determination of temperature stability was assayed using CD spectroscopy. Protein was dialyzed against 10 mM Na_2HPO_4, 5 mM NaCl, pH 7.4 overnight. CD spectra were recorded with a Jasco J-810 spectropolarimeter equipped with a Peltier-type temperature controller. The CD spectrum was recorded from 200 to 240 nm at different temperatures from 46 to 90 °C. A scanning speed of 50 nm/min and a heating rate of 1 K/min were used. Spectra were recorded every 1 K. Data was analyzed using Origin2017 (OriginLab).

Biolayer interferometry. A BLItz system (ForteBio) was used to determine protein-protein interactions. The ARG2 sensor was equilibrated in H_2O for 10 min. The sEGFR was immobilized on the sensor using EDC/NHS coupling chemistry. The following five steps were performed to immobilize the sEGFR on the sensor: initial baseline in H_2O (30 s), activation of the surface with 10 µM EDC/5 µM NHS (300 s), loading of sEGFR (1µg/ml in 10 mM sodium acetate, pH 3.5) (600 s), quenching with 1 M ethanolamine (300 s), regenerating with 50 mM NaOH and final equilibration with H_2O. A loaded sensor can be used for multiple interaction measurements. The scheme for measuring protein-protein interactions is described in the following: initial baseline in H_2O (30 s), association with varying analyte concentrations (120 s), dissociation in H_2O (180 s), regeneration with 50 mM NaOH (40 s) and final baseline in H_2O (120 s). The EGF-mCherry concentration of the initial solution was determined in three replicates using the NanoDrop2000c (Thermo Fisher Scientific). Concentrations of measured samples were estimated from the dilution series. As a negative control for sEGFR binding, bovine serum albumin (BSA) was tested, proving no association. Finally, sensorgrams were fitted globally to a 1:1 binding model by Blitz Pro 1.2.1.3. Curves and global fits were plotted using Origin2017 (OriginLab).

Mass spectrometry. A sample of protein (0.1 mg/ml) in PBS buffer was incubated with a final concentration of 10 mM tris(2-carboxyethyl)phosphine (TCEP) for 30 min at room temperature. After incubation the sample was acidified to a final concentration of 0.1 % formic acid (FA). A final volume of 50 µl of used for mass spectrometry measurements. A non-reduced sample was treated similarly without addition of TCEP. The samples were separated on a LaChrom Ultra HPLC system (VWR) using a reversed phase column (MAbPac RP 4µm 2.1 x 50 mm, Thermo Scientific). Eluents were 0.1% FA in water and 0.1% FA in acetonitrile. The mass spectrometric analysis was performed with a micrOTOF-Q II (Bruker Daltonik).

Cell culture. A431, MCF7 and MDA-MB-231 (DSMZ) were cultivated in Roswell Park Memorial Institute medium RPMI 1640 (D8758, Sigma Aldrich) supplemented with 10 % (v/v) fetal calf serum (FCS) and 1 % (v/v) penicillin/streptomycin (P/S) (P4333, Sigma Aldrich). Human Dermal Fibroblasts adult (HDFa), Hela, HEK293 and HEK293-EGFR cells were cultivated in Dulbecco's Modified Eagle Medium (DMEM) supplemented 10 % (v/v) FCS and 1 % (v/v) P/S.

Fluorescence microscopy. 5×10^4 cells were seeded in 24-well plates on poly-D-lysine coated coverslips. On the next day, media was removed and cells washed twice with 1× PBS. Cells were incubated at 37 °C with the desired protein concentration followed by three washing steps. All samples were fixed for 20 min with 4 % PFA (paraformaldehyde) and counterstained with 10 µM DAPI (4′,6-Diamidino-2-phenylindole) for 10 min. Finally, coverslips were mounted with Mowiol 4.88/Dabco (Roth). Images were acquired using a DMI6000B fluorescence microscope (Leica) with a 5× objective for wound healing assays (Figure 2), a 20× objective for EGFR expression level (Figure 3A), and a 63× objective with oil for EGF-mCherry internalization (Figure 3C). mCherry fluorescence was imaged with a TX2 filter bloc (ex 540-580, dc 595, em 607-683) and DAPI with a 405 filter bloc (ex 375-435, dc 455, em 445-495).

Wound healing assay. A431 were used to determine the biological activity of hEGF proteins in a wound healing assay. A431 cells (2×10^5 cells/ well) were seeded in 24-well plates (TC-treated, Sarstedt) and cultivated in RPMI-1640 culture media supplemented with 10 % FCS and 1 % P/S at 37 °C and 5 % CO_2 for 24 h until a confluent monolayer was formed. Cells were serum-starved overnight in RPMI-1640 without FCS. Two scratches per well were made using a sterile 10 µL pipet tip. After a washing using PBS, cells were treated with either the purified or a commercial hEGF (PHG0311, Gibco/Thermo Fisher Scientific), both at concentrations of 1 nM and 10 nM in RPMI-1640 with 1 % P/S. Control cells received an equal volume of buffer (1000 µL RPMI-1640 with PBS). Cells were maintained in cultivation conditions and cell migration was observed using a Leica DMI6000 microscope at selected time points ranging from 0 h to 8 h of incubation. The analysis of identical positions is guaranteed by the Mark and Find feature of Leica in combination with a high-precision motorized stage. The area of cell-free surfaces was determined using the MRI Wound Healing Tool plug-in (Volker Bäcker) for ImageJ version 1.51.[2] A linear regression of the surface area versus time was used to determine the cell migration rate in Origin2017(OriginLab) using $v_{migration} = \frac{|Slope|}{2 \times length\ of\ the\ gap}$. The length of the gap was measured using ImageJ.

Cell binding and competitions assay. For the cell-based fluorescence assay, 7×10^4 cells per well were seeded in 1 mL medium on a 24-well plate and incubated overnight at 37 °C. The medium was removed and cells were serum starved in 1 ml DMEM for 4 h at 37 °C. After washing with 1× PBS, 250 µl of the desired EGF-mCherry concentration in DMEM (w/o FCS and P/S) were applied to the cells ad incubated for 15 min at 37 °C. In competition assays EGF-mCherry (50 nM) and EGF (Gibco) (varying concentrations) were preincubated before application to cells. Cells were washed twice with PBS before detachment of the cells using 100 µl Accutase cell detachment solution for 10 min at 37 °C. Cells were diluted to a final volume of 200 µl and the cell suspension was transferred to a black 96-well plate (Sarstedt). Fluorescence measurements was carried out in duplicates using a Spark microplate reader (Tecan) set to 550 excitation and 610 nm emission. Data analysis was performed using Origin2017 (OriginLab) using the Dose response fit ($y = A1 + \frac{(A_2 - A_1)}{1 + 10^{(\log x_0 - x) \times p}}$). EC_{50} and IC_{50} values are given by the parameter $\log x_0$ and p represents the hill slope.

SDS-PAGE. Samples with 5× SDS loading buffer were prepared and incubated at 95 °C for 10 min followed by centrifugation for 1 min at 10,000×g and finally 20 µl per lane were loaded on a 12 % SDS-polyacrylamide gel (Hoefer SE260). Polyacrylamide gels were stained with Coomassie Blue staining solution.

■ REFERENCES

(1) Hsu, S.-C., and Hung, M.-C. (2007) Characterization of a novel tripartite nuclear localization sequence in the EGFR family. *J. Biol. Chem. 282*, 10432–40.

(2) Schindelin, J., Rueden, C. T., Hiner, M. C., and Eliceiri, K. W. (2017) The ImageJ ecosystem: An open platform for biomedical image analysis. *Mol. Reprod. Dev. 82*, 518–29.

■ SUPPLEMENTAL FIGURES

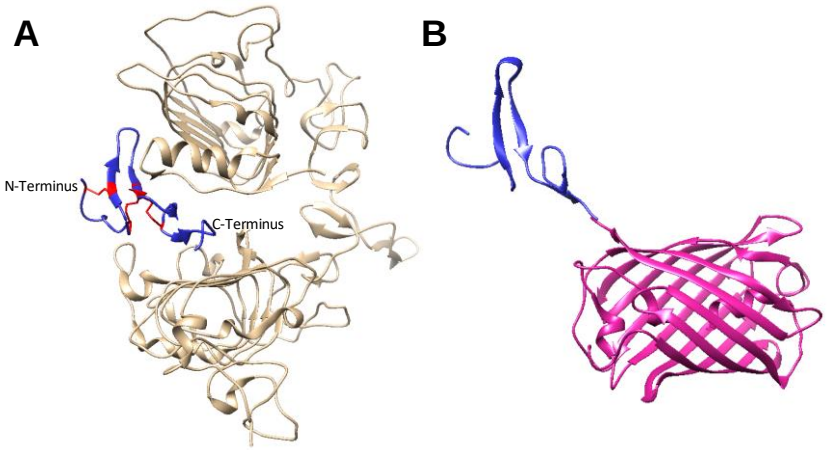
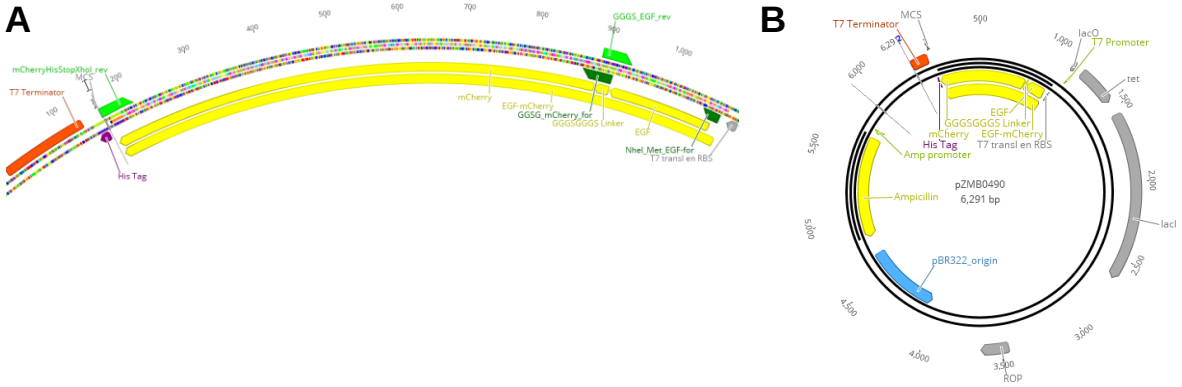


Figure S1. Crystal structure representation of EGF and EGF-mCherry. (A) EGF:EGFR crystal structure showing both termini to be amenable for fusion to other proteins (PDB ID 1IVO). Three essential disulfide bonds for EGF are highlighted in red. (B) Cartoon of EGF-mCherry from PDB ID 1JL9 and 2H5Q.



C EGF-mCherry with His-tag and restriction sites DNA sequence

GCTAGCATGAACAGCGACAGCGAGTGCCCCCTGAGCCACGACGGCTACTGCCTGCACGACGGCGTGTGCATGTACATCGAGGCCCTG
GACAAGTACGCCTGCAACTGCGTGGTGGGCTACATCGGCGAGAGGTGCCAGTACAGGGACCTGAAGTGGTGGGAGCTGAGGGGTGGT
GGATCTGGTGGAGGATCTATGGTGTCCAAGGGCGAAGAGGACAACATGGCCATCATCAAAGAGTTCATGCGGTTCAAGGTGCACATG
GAAGGCAGCGTGAACGGCCACGAGTTCGAGATTGAGGGCGAGGGCGAAGGCAGACCCTATGAGGGAACCCAGACCGCTAAGCTGAAA
GTGACCAAGGGCGGACCCCTGCCCTTCGCCTGGGATATCCTGTCCCCTCAGTTTATGTACGGCAGCAAGGCCTACGTGAAGCACCCC
GCCGACATCCCCGACTACCTGAAGCTGTCTTTCCCAGAGGGCTTCAAGTGGGAGAGAGTGATGAACTTCGAGGACGGCGGCGTCGTG
ACCGTGACCCAGGATAGCTCTCTCCAGGACGGCGAGTTTATCTACAAAGTGAAGCTGCGGGGCACCAACTTCCCTTCCGACGGCCCT
GTGATGCAGAAAAAGACAATGGGATGGGAGGCCAGCTCCGAGCGGATGTACCCTGAAGATGGCGCTCTGAAGGGCGAGATCAAGCAG
CGGCTGAAGCTGAAGGATGGCGGCCACTACGACGCCGAAGTGAAAACCACCTACAAGGCCAAAAAGCCCGTGCAGCTGCCTGGCGCT
TACAATGTGAACATCAAGCTGGACATCACCTCCCACAACGAGGACTACACCATCGTGGAACAGTATGAGCGGGCCGAGGGCAGACAT
TCCACCGGCGGCATGGATGAGCTGTACAAGCACCACCACCACCACCACTGA**CTCGAG**

EGF-mCherry with His-tag protein sequence

MASMNSDSECPLSHDGYCLHDGVCMYIEALDKYACNCVVGYIGERCQYRDLKWWELRGGGSGGGSMVSKGEEDNMAIIKEFMRFKVH
MEGSVNGHEFEIEGEGEGRPYEGTQTAKLKVTKGGPLPFAWDILSPQFMYGSKAYVKHPADIPDYLKLSFPEGFKWERVMNFEDGGV
VTVTQDSSLQDGEFIYKVKLRGTNFPSDGPVMQKKTMGWEASSERMYPEDGALKGEIKQRLKLKDGGHYDAEVKTTYKAKKPVQLPG
AYNVNIKLDITSHNEDYTIVEQYERAEGRHSTGGMDELYKHHHHHH

Figure S2. Schematic overview of pET21a(+)-EGF-mCherry. (A) Plasmid map section of EGF-mCherry from pET21a. Primers used in PCR reactions are highlighted in light or dark green with their distinct name. (B) Plasmid map of pET21a(+)_EGF-mCherry. The fusion construct was integrated into the backbone via NheI and XhoI restriction enzymes. (C) DNA and protein sequences of EGF-mCherry, flanking XhoI and NheI sites are underlined in the DNA sequence.

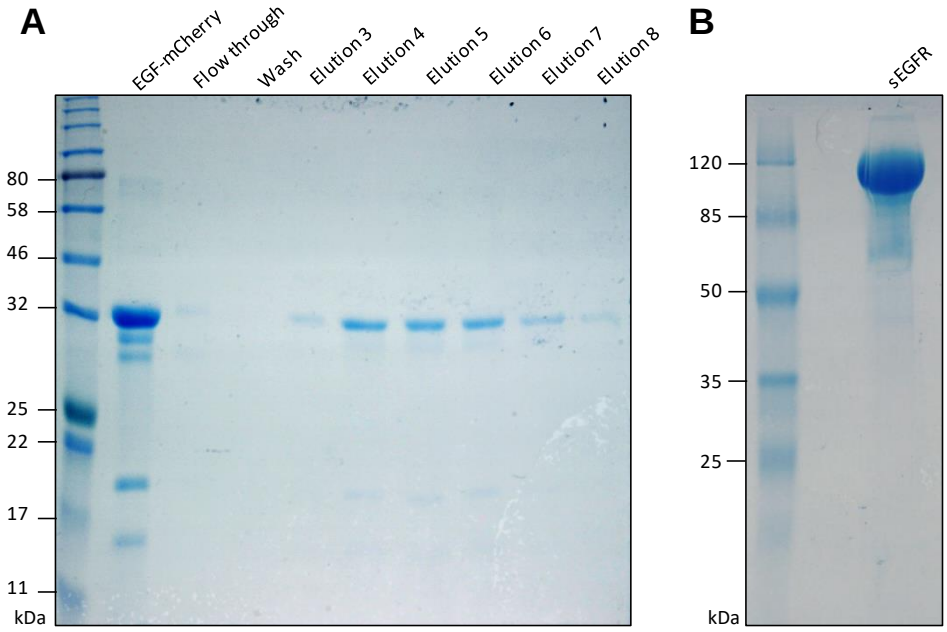


Figure S3. SDS PAGE of EGF-mCherry and sEGFR after purification via IMAC. (A) Visualization of EGF-mCherry after IMAC and following anion exchange chromatography. (B) sEGFR was secreted from 293F suspension culture after transient transfection. The soluble receptor was purified via IMAC.

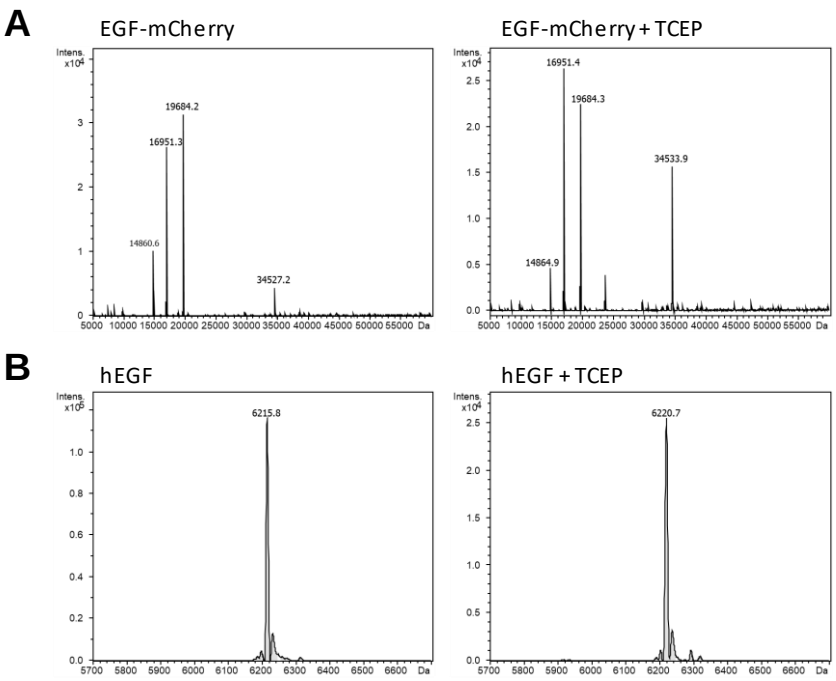


Figure S4. LC-ESI-MS analysis of EGF-mCherry and hEGF (Gibco™). (A) Visualization of the whole spectrum range for EGF-mCherry incubated either without or with TCEP. The theoretical mass for the primary sequence of EGFmCherry was calculated with 34,555 Da. During cyclization of the chromophore tripeptide water is eliminated and in a second step the N-acylimine is formed leading to a deviation in mass of -22 Da. Oxidation of the free thiols (-6 Da) results in the expected mass of 34,527 Da, which was found (left panel). The mass peak of 16,951 Da can be attributed to the internal standard. Two masses of 14,860 Da (N-terminal part EGF-mCherry) and 19,684 Da (C-terminal part of EGF-mCherry) can be explained by the degradation products of EGF-mCherry. LC-MS is performed on a column heated to 65 °C, which led to the partial degradation of the N-acylimine. Under reducing conditions the N-acylimine seems more stable. A more detailed overview is given in Figure S5C. (B) Spectrum for hEGF (Gibco™) without and with TCEP.

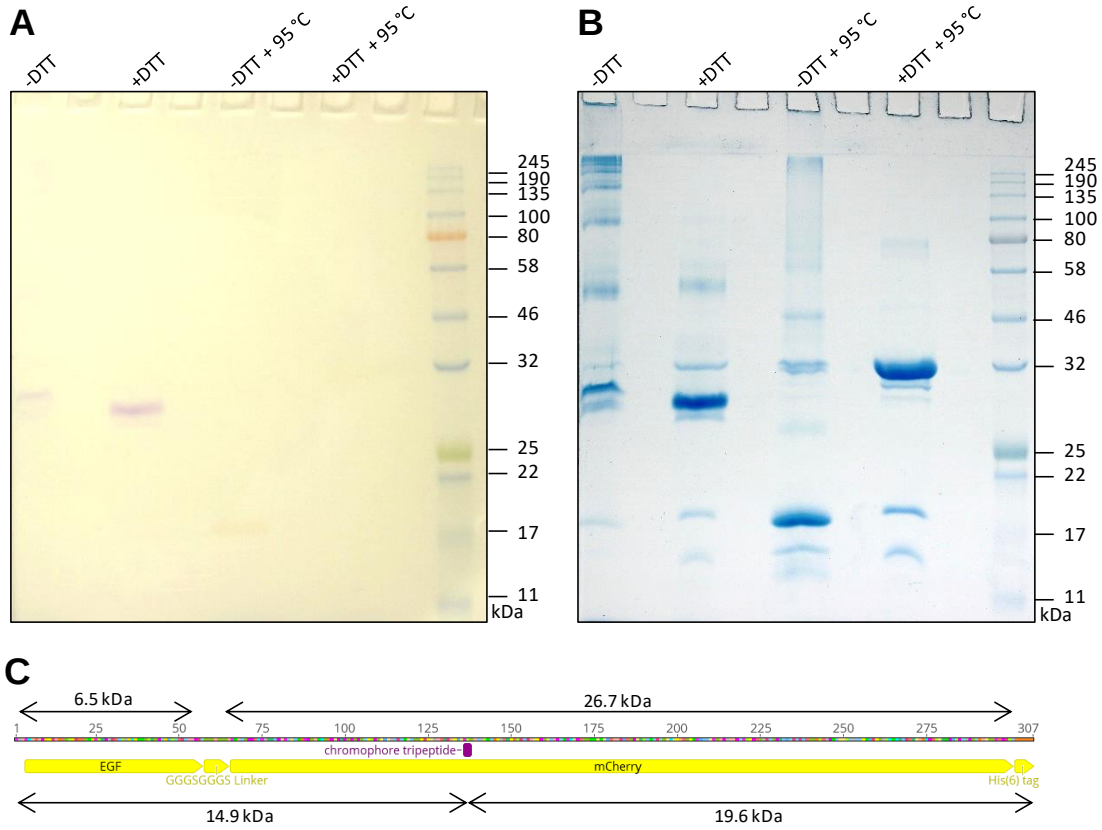


Figure S5. SDS PAGE of EGF-mCherry under various conditions. EGF-mCherry was incubated with reducing (+DTT) and without reducing agent (–DTT) either at room temperature or at 95 °C for 10 min before application to a 12 % SDS-PAGE. Images under white light are shown before (A) and after staining (B) with Coomassie Blue. Without staining the prestained marker and the functional mCherry chromophore are visible. (C) Schematic EGF-mCherry fusion. EGF has a size of 6.5 kDa and mCherry of 26.7 kDa. The overall mass of 34.5 kDa corresponds to the whole fusion construct with glycine-serine linker and His_6-tag. The chromophore tripeptide highlights the position of the N-acylimine, which is preferentially hydrolyzed upon heat denaturation. Two fragments result from this hydrolysis corresponding to 14.9 kDa and 19.6 kDa.

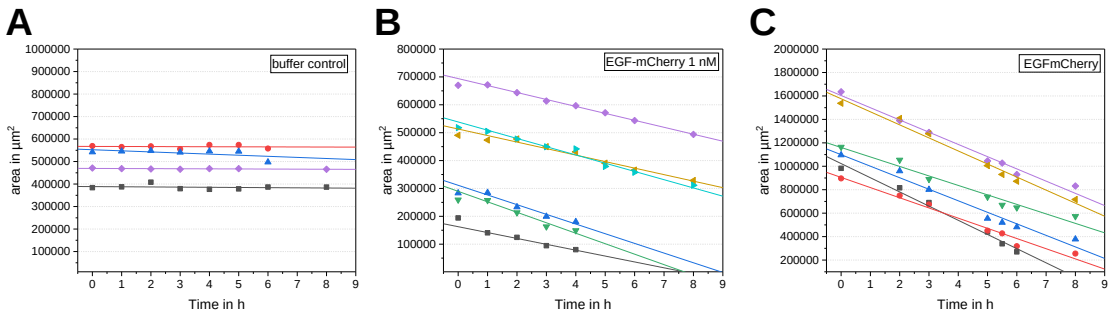


Figure S6. Wound healing assay. Analysis of the wound area during the course of the experiment for the buffer control (A), EGF-mCherry (B) and hEGF (Gibco™) (C). Linear regressions for all replicates are shown.

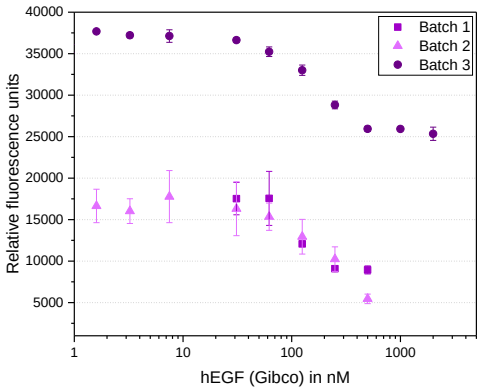


Figure S7. EGF-mCherry competition assay. Different batches of EGF-mCherry production were used for competition assays with hEGF (Gibco™). Differences in mean fluorescence were expected as these experiments were independently performed with different gains at the Tecan Spark. Batch 3 represent the batch from Figure 4B. The IC_{50} values were not explicitly determined, but the curves of the three batches are similar, indicating reproducibility of the experiment.