

RESEARCH

Open Access



# Expression of acid cleavable Asp-Pro linked multimeric AFP peptide in *E. coli*

Murad Mollaev<sup>1,2,3</sup>, Artur Zabolotskii<sup>3,4</sup>, Neonila Gorokhovets<sup>5</sup>, Elena Nikolskaya<sup>3,6</sup>, Maria Sokol<sup>3,6</sup>, Andrey Tsedilin<sup>7</sup>, Mariia Mollaeva<sup>3,6</sup>, Margarita Chirkina<sup>3,6</sup>, Timofey Kuvaev<sup>8</sup>, Anna Pshenichnikova<sup>1</sup> and Nikita Yabbarov<sup>3,6\*</sup>

## Abstract

**Background:** Difficult to express peptides are usually produced by co-expression with fusion partners. In this case, a significant mass part of the recombinant product falls on the subsequently removed fusion partner. On the other hand, multimerization of peptides is known to improve its proteolytic stability in *E. coli* due to the inclusion of body formation, which is sequence specific. Thereby, the peptide itself may serve as a fusion partner and one may produce more than one mole of the desired product per mole of fusion protein. This paper proposes a method for multimeric production of a human alpha-fetoprotein fragment with optimized multimer design and processing. This fragment may further find its application in the cytotoxic drug delivery field or as an inhibitor of endogenous alpha-fetoprotein.

**Results:** Multimerization of the extended alpha-fetoprotein receptor-binding peptide improved its stability in *E. coli*, and pentamer was found to be the largest stable with the highest expression level. As high as 10 aspartate-proline bonds used to separate peptide repeats were easily hydrolyzed in optimized formic acid-based conditions with 100% multimer conversion. The major product was represented by unaltered functional alpha-fetoprotein fragment while most side-products were its formyl-Pro, formyl-Tyr, and formyl-Lys derivatives. Single-step semi-preparative RP-HPLC was enough to separate unaltered peptide from the hydrolysis mixture.

**Conclusions:** A recombinant peptide derived from human alpha-fetoprotein can be produced via multimerization with subsequent formic acid hydrolysis and RP-HPLC purification. The reported procedure is characterized by the lower reagent cost in comparison with enzymatic hydrolysis of peptide fusions and solid-phase synthesis. This method may be adopted for different peptide expression, especially with low amino and hydroxy side chain content.

**Keywords:** Multimer expression, Asp-Pro cleavage, Alpha-fetoprotein, Recombinant peptide

## Background

Recombinant peptide expression is an alternative to solid-phase synthesis. Relatively short polypeptides are difficult to express in *E. coli*, which is probably due to the active proteolysis of cytosolic soluble products [1, 2]. The common way to improve the stability of peptides in the cytosol is to co-express peptides with fusion partners such as PurF fragment, ketosteroid isomerase,

PaP3, and TAF12 domain. Moreover, the fusion partner can be the same peptide itself, when expressed in tandem repeats [3–5]. Multimeric expression as well as fusion co-expression requires the specific enzymatic or chemical cleavage site(s). Enzymatic digestion with factor Xa (IEGR↓), enterokinase (DDDDK↓), thrombin (LVPR↓GS), TEV protease (ENLYFQ↓G), and HRV 3C protease (LEVLFG↓GP) is the most common methods for hydrolysis of fusion proteins [6–9]. These methods are highly specific, but limited to enzyme costs and sensitivity to environmental changes. Denaturing conditions can be detrimental to enzyme activity, while most

\*Correspondence: marvint@inbox.ru

<sup>3</sup> JSC Russian Research Center for Molecular Diagnostics and Therapy, 8 Simferopolsky boulevard, Moscow 117638, Russia  
Full list of author information is available at the end of the article

peptide multimers tend to form aggregates. Alternatively, chemical cleavage has lower reagent costs, wider temperature and suitable pH range, solubilization agent compatibility, and shorter artificial sequence insertions [10, 11]. Some peptide bonds are known to be less acid stable than others. Acid lability of  $-D-X-$  bond (where X = any amino acid) was discovered by Partridge and Davis [12], but the technique was of little use due to the extremely critical conditions of such cleavage. Among  $-D-X-$  bonds,  $-D-P-$  and  $-D-C-$  were found to be the most acid sensitive [13, 14]. The mechanism of  $-D-P-$  bond acid hydrolysis is that the N atom of proline attacks protonated sidechain carboxyl of aspartate, thus forming an unstable cation-imide intermediate, which is then become rapidly hydrolyzed. Due to the lower abundance of  $-D-P-$  bonds, this technique has an advantage over BrCN cleavage [15]. The use of such expression system has great prospects in the production of low molecular weight peptides for targeted therapy (highly specific inhibitors or agonists, vector molecules, etc.), since the low molecular weight peptides tend to be more proteolytically stable in multimers and have a low frequency of  $-D-P-$  bonds. Human AFP is an albumin family protein that primarily functions as a transporter of lipophilic molecules and several ions during the fetal period. AFP is capable of triggering various signaling pathways, including those that promote immunosuppression [16]. The expression level of AFP in healthy adults is significantly lower than in the fetal period; however, AFP is detected in patients with hepatocellular carcinoma and some other cancers [17]. AFP is known to selectively bind and to be internalized by a wide range of cancer cells. Different researchers reported its suitability for anticancer drug delivery systems as targeting motif [18–21]. Recent studies showed that endogenous full-length AFP may promote hepatic cancer progression and is not recommended for human treatment; however, the mechanism is not completely understood [22, 23]. Shorter functional fragments may still prove its safety [22], while remaining specific activity. Arguably, short fragments may block epitopes recognized by native endogenous AFP, thereby inhibiting its immunosuppressive and tumor-stimulating functions. Previously, the KQEFLIN peptide was found to be the minimal and necessary AFP part for receptor binding [24]. Here, we report the multimeric  $-D-P-$  linked extended AFP receptor-binding fragment expression with optimized multimerization, cleavage, and purification conditions.

## Methods

The reagents used in this study are provided in Supplementary information 2.

## Cloning and expression

Molecular cloning was performed by the standard procedures [25]. Plasmid DNA was extracted using the Quantum Prep Plasmid Miniprep Kit. In order to amplify the sequence coding, PWGVALQTMKQEFLINLVKQKPQITD peptide A-C primers have been selected according to the human AFP mRNA (GenBank NM\_001134): (A) ATTCCATGGCTGATCCGTGGGGTGTAGCG (NcoI restriction site is underlined); (B) ATTGTCGACGATCCGTGGGGTGTAGCGC (Sall restriction site is underlined); (C) ATATCTCGAGCGGATCTGTAATTTGTGGC (XhoI restriction site is underlined). Insertion 1 (supplementary information Fig. S1) was amplified with the forward primer A and reverse primer C and with the previously designed plasmid, encoding C-end domain of AFP [26] as a matrix. Insertion 2 (supplementary information Fig. S1) was prepared in the same way by using primers B and C. Insertion 1 was cloned into pET-28a(+) expression vector using NcoI and XhoI. The resulting plasmid pET28AFPpep1 encoded monomeric peptide. To obtain plasmids encoding different numbers of tandem repeats (from 2 to 14), Insertion 2 was digested with Sall and XhoI and sequentially cloned into the XhoI site of pET28AFPpep1. The number of inserts was analyzed by PCR with T7 promoter and terminal primers followed by agarose electrophoresis. To ensure the forward direction of all inserts, PCR with T7 terminal and C reverse primers was used.

Constructed pET28AFPpepN plasmids were transfected into *E. coli* BL21(DE3) expression strain (NEB, USA). Clones were selected with kanamycin (50  $\mu$ g/ml) on agarized LB plates. One colony for each sample was inoculated into a 5-ml LB medium with kanamycin (50  $\mu$ g/ml) and cultured overnight at 37°C with shaking. Further cells were inoculated into fresh LB medium in ratio 1:50 vol. and cultured at 37°C with shaking until OD600 reached 0.8. The expression of pET28AFPpepN was induced by 1 mM IPTG for 4 h. For expression analysis, 50- $\mu$ l samples were taken at each stage and analyzed by SDS-PAGE.

## Isolation and purification of multimeric peptide

Multimeric peptide AFPpep5 (MADPWGVALQTMKQEFLINLVKQKPQITDPLDDPWGVALQTMKQEFLINLVKQKPQITDPLDDPWGVALQTMKQEFLINLVKQKPQITDPLDDPWGVALQTMKQEFLINLVKQKPQITDPLEHHHHH) was purified by immobilized metal affinity chromatography on Ni<sup>2+</sup> Sepharose 6 Fast Flow (GE Healthcare). For this purpose, 40 ml of Ni<sup>2+</sup> Sepharose 6 Fast Flow was packed into a 50  $\times$  150 column and equilibrated with binding buffer (50 mM Tris-HCl

pH 7.5, 6 M guanidine hydrochloride). The wet biomass was resuspended in 100 ml of 50 mM Tris-HCl buffer (pH 7.5) on ice and sonicated 3 times, 40 strokes each. The suspension was centrifuged at 10,000g, 4°C for 15 min. The precipitate containing inclusion bodies was thrice resuspended in the same buffer and centrifuged. Washed precipitate was dissolved in 100 ml of 50 mM Tris-HCl buffer (pH 7.5) containing 6 M guanidine hydrochloride and centrifuged at 10,000 rpm (4 °C, 15 min). Solubilized inclusion bodies (6 g) from the supernatant were applied to the column, followed by washing with 3 volumes of binding buffer. Multimer was eluted by 1 volume of elution buffer (50 mM Tris/HCl pH 7.5, 6 M guanidine hydrochloride, 0.3 M imidazole), followed by dialysis against milli-Q water at 4°C. The precipitated multimeric peptide was centrifuged at 8000g, 4°C for 10 min. Sedimented pellets were resuspended in milli-Q water and lyophilized. Samples after each purification step were taken and analyzed by SDS-PAGE. The identity and purity of lyophilized multimeric peptide were determined by MALDI-MS, SDS-PAGE, and RP-HPLC.

#### Cleavage of multimeric peptide to AFPpep1 by acidic hydrolysis

To cleave –D–P– bond between AFPpep monomers, several methods were applied [27]. Multimer was dissolved in corresponding hydrolysis buffer (A-L) (Table 1) and left for stirring. Every step sample from the reaction mixture was taken, neutralized by 1 M NH<sub>4</sub>OH, and frozen at –20°C for hydrolysis termination. Then, samples were analyzed by tricine SDS-PAGE.

#### Purification of AFPpep1 by RP-HPLC

The hydrolysis mixture (D, 72h) was neutralized with 1 M NH<sub>4</sub>OH to pH 7.0. Peptide sediment was collected

by centrifugation at 10,000g for 10 min. Pellets were resuspended in deionized water and lyophilized. The lyophilized mixture was diluted with mobile phase consisted of 0.1% TFA in water (mobile phase A) and pure AcN (mobile phase B) in a ratio of 70:30, v/v. AFPpep1 was purified using 1525 binary pump and 2487 UV–VIS detector (Waters, USA). Separation was performed on Symmetry Prep C18 column with the dimensions of 300 mm × 7.8 mm ID × 7 μm. Isocratic elution at 30% B for 30 min followed by a linear gradient from 30 to 40% B for 20 min and 40 to 70% B for 10 min. The flow rate was kept constant at 3 ml/min. Peaks were detected at 214 nm. The major product peaks were collected, lyophilized, and analyzed by MALDI-MS.

#### MALDI-MS

MALDI-MS analysis was performed on UltrafleXtreme TOF/TOF high-resolution mass spectrometer (Bruker Daltonik GmbH, Germany) equipped with Smartbeam II UV laser with following conditions: 10 mg/ml 2,5-dihydroxybenzoic acid solution (20% aqueous AcN, 0.5% TFA) as matrix (0.5 μl per 1.5 μl of the sample), MALDI source in positive mode, analyzer in reflectron mode, and scan range *m/z* 500–6500. Spectra were processed with FlexAnalysis 3.3 (Bruker Daltonik GmbH, Germany). The detected *m/z* of peptides were compared with the theoretically calculated in accordance with the known amino acid sequences of AFP (NP - 001125.1) and the known genetic sequences of constructs.

#### LC-MS

HPLC-ESI-MS analysis was performed on Impact II QqTOF high-resolution mass spectrometer (Bruker Daltonik, Germany) with Elute UHPLC (Bruker Daltonik, Germany) on Intensity Solo 1.8 C18-2 2.1 × 100 mm 1.8

**Table 1** Hydrolysis conditions for multimeric peptide cleavage

Method	Hydrolysis conditions	Time, h
A	10% acetic acid, adjusted to pH 2.5 with pyridine, 37°C	48–96
B	10% acetic acid, 25% propanol-1 adjusted to pH 2.5 with pyridine, 37°C	48–96
C	50% formic acid, 37°C	48–96
D	70% formic acid, 37°C	24–96
E	90% formic acid, 37°C	48–72
F	0.1 M HCl + 12%SDS, 37°C	24–120
G	10% acetic acid + 12%SDS, 37°C	24–120
H	0.1 M HCl + 12%SDS, 95°C	0.5–10
I	10% acetic acid + 12%SDS, 95°C	0.5–10
J	Tris-HCl 50 mM pH 7.0 (at RT) + 12%SDS, 95°C	0.5–10
K	10% acetic acid, adjusted to pH 2.5 with pyridine + 6M guanidine HCl, 37°C	24–72
L	0.1 M HCL + 6M guanidine HCl, 37°C	24–72

$\mu\text{m}$  90 Å reverse-phase column (Bruker Daltonik, Germany) with the following conditions: column flow 0.25 ml/min, gradient elution from 5 to 70% B in 25 min (A: 0.1% formic acid in water, B: 0.1% formic acid in AcN), column temperature 40°C, injection volume 5  $\mu\text{l}$ , ESI source in positive mode, HV capillary at 4.5 kV, spray gas–nitrogen at 2.1 bar, dry gas–nitrogen at 8 l/min 220°C, scan range  $m/z$  50–2200, 2-Hz scan rate for full scan, automatic MS/MS mode (CID) with dynamic scan rate 2–8 Hz, nitrogen as collision gas, collision energy from 23 eV at  $m/z$  300 to 65 eV at  $m/z$  >1300, and automatic internal calibration with ESI-L low concentration tuning mix (Agilent Technologies, USA). Spectra were processed with BioPharma Compass 3.1.1 (Bruker Daltonik, Germany).

#### Cell binding and internalization inhibition assay

Human mammary gland adenocarcinoma cells (MCF-7) were maintained in DMEM medium supplemented with 10% fetal bovine serum and gentamycin (50  $\mu\text{g}/\text{ml}$ ) in a CO<sub>2</sub> incubator at 37°C in a humidified atmosphere containing 5% CO<sub>2</sub>. Cells were replated with Trypsin-EDTA solution twice per week. Before the experiment, cells were incubated for 2 h in serum-free DMEM. The concurrent binding and internalization with fluorescein-labeled AFP 3rd domain (3dAFP-FITC) was accessed in order to analyze the AFP<sub>pep1</sub> functional activity. The cells were incubated in the presence of 18.5  $\mu\text{M}$  3dAFP-FITC during 1 h at 4°C or 37°C. AFP<sub>pep1</sub> (0.182 mM or 0.281 mM) and not labeled 3dAFP (18.5  $\mu\text{M}$  or 37  $\mu\text{M}$ ) were used to interfere with 3dAFP-FITC interaction with AFP receptors. Different concentrations were used to confirm that the inhibition is dose dependent in chosen diapason. A self-inhibition test of 3dAFP-FITC with unlabeled 3dAFP-FITC was used as a control to confirm

that this assay works properly. Different temperatures were chosen to find out whether AFP<sub>pep1</sub> may inhibit binding to the cell surface or active endocytosis which is in part restricted at +4°C. The primary stock solutions of peptides were prepared in DMSO followed by PBS dilution; low-intensity ultrasound was applied if necessary. The fluorescence intensity of cells was measured with a Dako CyAn ADP 9 flow cytometer equipped with an argon laser (488 nm, 525 nm FITC band-pass). For each sample ( $2 \times 10^3$  cells), the median fluorescence intensity (MFI) was determined. Unstained cells were used as control.

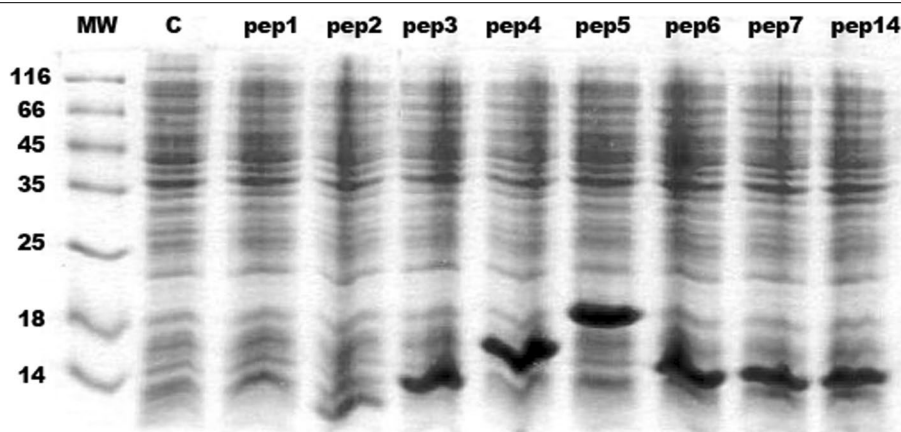
## Results

### Pentameric peptide expression improves yield

Recombinant expression of short polypeptides tends to be complicated mostly due to the low proteolytic stability in *E. coli*. All our tries to express monomeric AFP fragment PWGVALQTMKQEFLINLVKQKPQITD (AFP<sub>pep1</sub>) were unsuccessful. Multimerization facilitated aggregation of the peptide into inclusion bodies. The expression level was proportional to the degree of multimerization, but increased only to 5 tandem repeats (Fig. 1). Transformation with higher than 5 multimeric sequences led to the expression of major product peptide with MW close to pep3 in all cases, probably due to the low intracellular stability of repeated constructs [5]. For further experiments, strain transformed with pentamer (AFP<sub>pep5</sub>) was used because of the highest stable productivity.

### Purification and hydrolysis of multimer

For purification purposes, a cleavable his-tag sequence was inserted at C-end. Standard Ni<sup>2+</sup> chromatography followed by low ionic strength precipitation led to



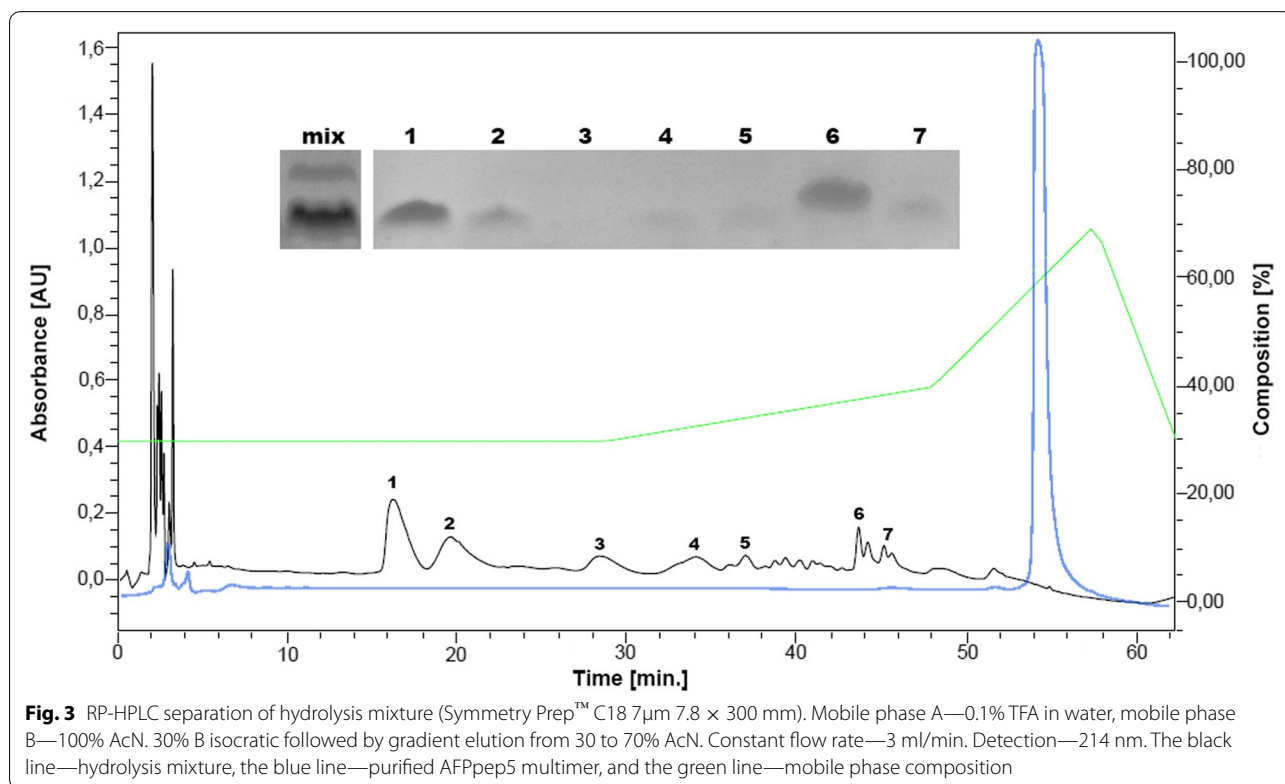
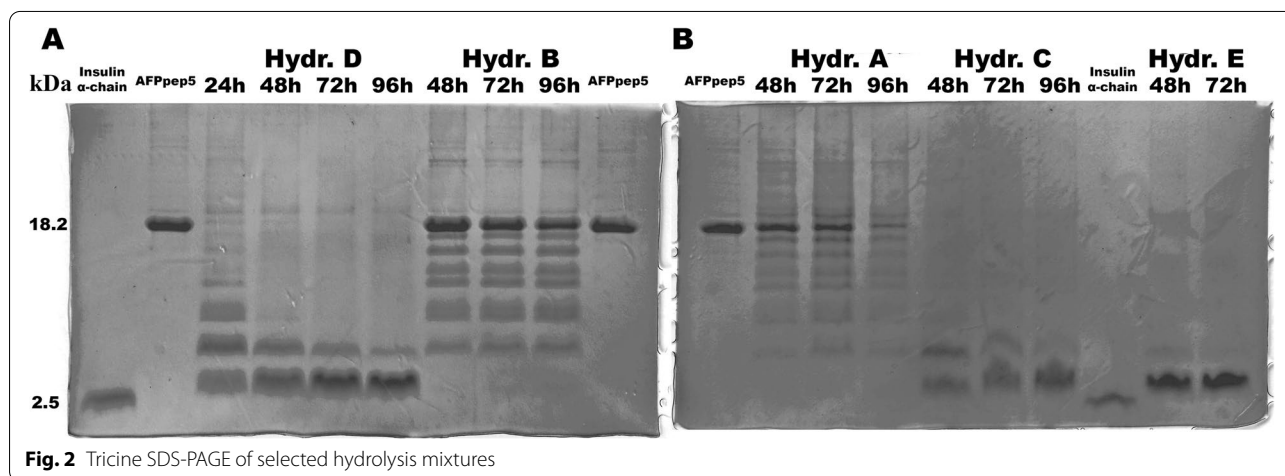
**Fig. 1** SDS-PAGE of IPTG induced *E. coli* BL21 (DE3) transformed with 1–7 and 14 gene repeats (pep1–pep7, pep14), biomass samples. Control—untransformed cells. Bacteria transformed with 8–13 repeats behaved similar to pep7 and pep14 samples on SDS-PAGE (unpublished data)

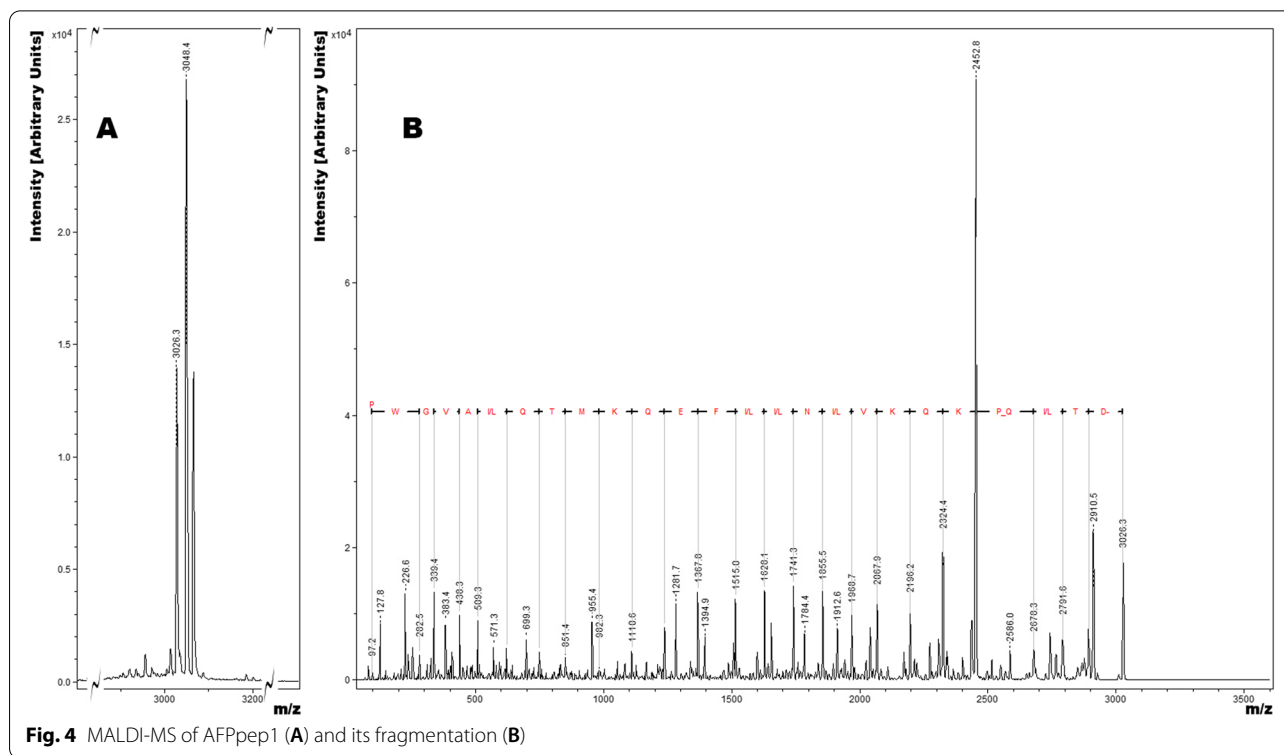


approximately 75% yield of purified pentamer. All of the peptide repeats and his-tag were divided with  $-D-P-$  linker. Several mild acidic conditions listed in Table 1 are tested to selectively hydrolyze  $-D-P-$  bonds. Based on tricine SDS-PAGE (selected samples, Fig. 2; other, Supplementary information Fig. S2), optimal cleavage with close to 100% conversion and selective formation of  $\approx 3$  kDa fraction was with either 50–90% formic acid or in SDS-containing tris and HCl solutions.

**Monomer purification and characterization**

AFPpep1 was purified from the formic acid hydrolysis mixture by semi-preparative RP-HPLC. Despite the fact that the majority of peptides in the hydrolysis mixture had similar molecular weight (Fig. 2, Supplementary information Fig. S3), 7 main fractions with different retention times were isolated (Fig. 3). The congruence of peak 1 (Fig. 3) to the AFPpep1 sequence was confirmed by MALDI-MS analysis (Fig. 4). On the first spectrum (Fig. 4A), only 3 peaks were detected, corresponding to





**Fig. 4** MALDI-MS of AFPpep1 (A) and its fragmentation (B)

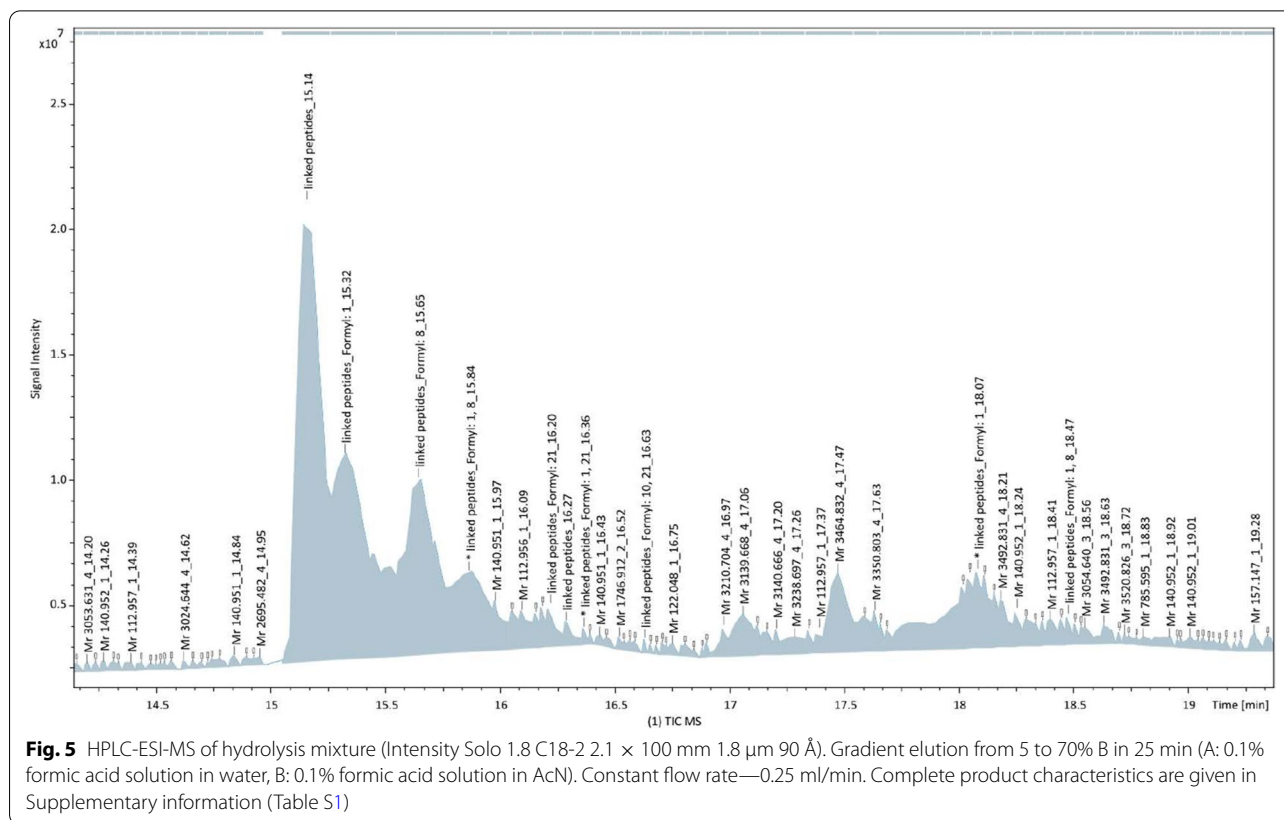
AFPpep1 (3026), its Na<sup>+</sup> (+22Da), and K<sup>+</sup> (+38Da) salts. On the fragmentation spectrum (Fig. 4B), the declared sequence was completely assembled. The closest difficult-to-separate impurity was the monoformylated (+28 Da) derivative AFPpep1 (peak 2, Fig. 3). Dimer, its mono- and di-formylated derivatives, and products of incomplete hydrolysis (AFPpep1-PLDD, AFPpep1-PLHHHHHH), as well as products of less specific hydrolysis at W–G, T–D, and Q–T bonds, were detected in the reaction mixture but not in the first two peaks. MALDI-MS failed to obtain the fragmentation spectrum of the formylated derivatives of AFPpep1. However, to the best of our knowledge, previously not detected direct formylation at the N-terminal Pro was observed in the cleaved fragment of the C-end containing affinity tag - PLEHHHHH (Supplementary information Fig. S4). Subsequently, the formylated derivatives of AFPpep1 were determined by HPLC-ESI-MS (Fig. 5). The most significant formylation sites were found to be the N-terminal Pro, similar to the PLEHHHHH fragment, and Tyr8 (O-formylation). N-formylation at Lys10, Lys19, and Lys21 was also detected, but to a much lesser extent. This selectivity of formylation may be the result of kinetic or steric differences. It is possible that formylation of the imino group of N-terminal proline may be specifically coupled to cleavage reaction.

#### Cell binding and internalization

Cell binding and internalization ability of purified AFPpep1 was accessed through inhibition of complete AFP 3rd domain (3dAFP) described previously [26]. Self-inhibition of cell binding and internalization with labeled and unlabeled 3dAFP was performed as a control (Fig. 6A, B). Co-incubation of AFPpep1 with labeled 3dAFP led to a decrease of fluorescent signal due to concurrent binding (Fig. 6C) and internalization (Fig. 6D). Comparable inhibition of cell binding and internalization was observed only with a high molar excess of the peptide, which is probably due to the greater affinity of the large protein fragment. On the other hand, limited inhibition may result from more specific binding of the peptide to the canonical AFP receptor.

#### Discussion

Sequential multimerization of the AFP peptide fragment from 1 to 5 repeats led to a gradual increase in the stability of the expression product, while multimers containing 6 or more repeats were not detected by SDS-PAGE. Thus, the expression of multimers tends to be possible up to a certain number of repeats, which is a sequence-specific value. For the selected AFP fragment, the pentamer turned out to be the largest in size among

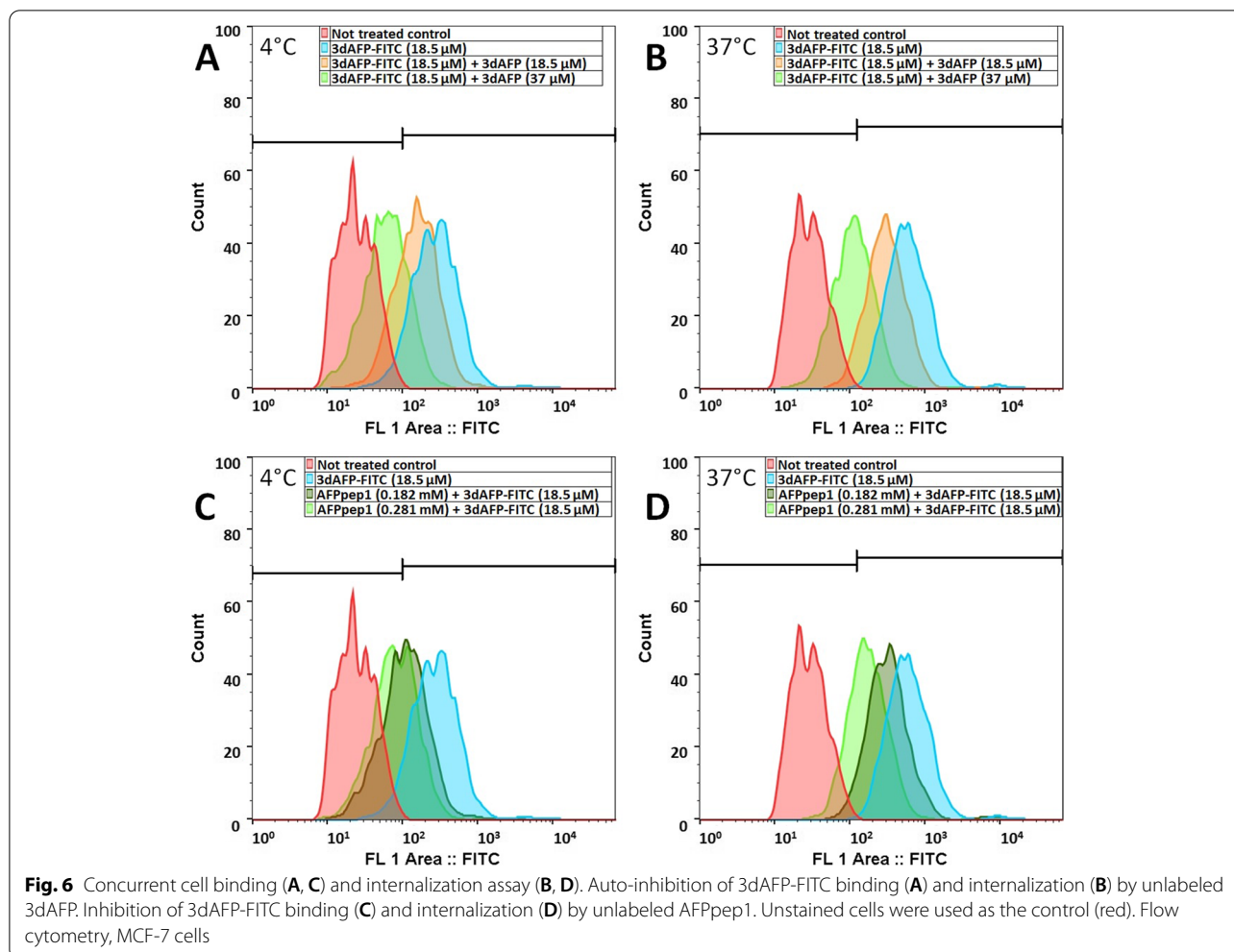


the stable ones and had the highest level of expression. Various methods of the multimer hydrolysis at  $-D-P-$  bonds were analyzed; however, we were not lucky to overperform the known standard method with formic acid. Multimer solubility in hydrolysis buffer seems to be the most critical parameter for optimal  $-D-P-$  cleavage while formic acid is best suited for dissolving almost any peptide. Despite the fact that SDS-containing hydrolysis samples had bands with characteristic MW of hydrolysis fractions, we did not find any possible way to completely separate peptides from SDS, which may interfere with mass spectrometry. MALDI-MS confirmed the presence of monomeric AFP<sub>pep1</sub> (Supplementary information Fig. S3) only in formic acid cleaved mixtures (D, G, and J conditions were tested). We assume that if a method for separation of peptides from SDS was available, tris-buffered hydrolysis with SDS would provide a convenient alternative to formic acid. Hydrolysis in tris buffer can be easily controlled by temperature (as the temperature rises, pH shifts to the acidic) and does not require neutralization for termination. Also, side reactions with the formation of formyl derivatives would be excluded. After hydrolysis with formic acid, we managed to find a simple one-step semi-preparative RP-HPLC method for

separating the target peptide from formylated and less specific hydrolysis products. It is important to note that of all the formylation products, only Lys10 is part of Moro's minimal receptor-binding peptide [28]. The low degree of formylation at this site and the possibility of preparative separation (Fig. 3) will minimize the risks of loss of the receptor-binding activity of this peptide. The purified peptide structure was fully assembled by tandem mass spectrometry (Fig. 4), and the activity was confirmed by competitive inhibition of binding and internalization of the larger fluorescently labeled AFP fragment (Fig. 6).

## Conclusions

Multimeric expression of peptides can be performed with  $-D-P-$  bond linkers. Such technique reveals a high peptide expression level with localization in inclusion bodies, which supports cleavable purification tags insert. Gradually, with an increase in multimerization, the number of expressed repeats stabilizes; in our case, it stopped at 5. Probably, larger fragments are also expressed in insignificant undetectable amounts, but we did not find any evidence for that [29]. Cleavage of as high as 10 linkers in one construct can be done with up to 100% conversion and high selectivity using optimized conditions



in formic acid. The target monomeric product was represented by the largest peak on both analytical and semi-preparative HPLC and can be separated from close formyl-Pro, formyl-Tyr, and formyl-Lys derivatives. New formic acid-free hydrolysis methods still could prove even higher selectivity if other easily removable solubilization agents were found. Perhaps, development of selective deformylation methods could also contribute to the improvement of this technique. In this work, a method is described that allows one to obtain a recombinant AFP fragment in semi-preparative amounts in the highest quality and quantity and in the very efficient way. This fragment can find application in cytotoxic drug delivery [18, 24, 30] or as an inhibitor of endogenous AFP; however, more research is still to be done. The reported procedure is characterized by the lower reagent cost in comparison with enzymatic hydrolysis of peptide

fusions or solid-phase synthesis and may be adopted for different peptide expression, especially with low amino and hydroxy side chain content.

#### Abbreviations

(U)HPLC: (Ultra) high-performance liquid chromatography; 3dAFP: Alpha-fetoprotein (human) third (C-end) domain; AcN: Acetonitrile; AFP: Alpha-fetoprotein (human); AFPpep1: PWGVALQTMKQEFLLINLVKQKPKQITD; AFPpep5: MADPWGVALQTMKQEFLLINLVKQKPKQITDPLDDPWGVALQTMKQEFLLINLVKQKPKQITDPLDDPWGVALQTMKQEFLLINLVKQKPKQITDPLDDPWGVALQTMKQEFLLINLVKQKPKQITDPLDDPWGVALQTMKQEFLLINLVKQKPKQITDPLEHHHHH; BCA: Bicinchoninic acid; CID: Collision-induced dissociation; DMEM: Dulbecco's modified Eagle medium; DMSO: Dimethylsulfoxide; EDTA: Ethylenediaminetetraacetic acid; ESI: Electrospray ionization; FBS: Fetal bovine serum; FITC: Fluorescein isothiocyanate; HV: High voltage; IPTG: Isopropylthio- $\beta$ -galactoside; LB: Luria-Bertani medium; MALDI: Matrix-assisted laser desorption ionization; MCF-7: Human mammary gland adenocarcinoma cells; MFI: Mean fluorescence intensity; MS: Mass spectrometry; PAGE: Polyacrylamide gel electrophoresis; PCR: Polymerase chain reaction; RP: Reverse phase; SDS: Sodium dodecyl sulfate; TFA: Trifluoroacetic acid; TOF: Time of flight; UV-VIS: Ultraviolet-visible absorption.



## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s43141-021-00265-5>.

**Additional file 1:** Supplementary information. **Figure S1.** Cloning strategy. **Figure S2.** Tricine-SDS-PAGE analysis of different hydrolysis methods. **Figure S3.** Analysis of hydrolysis mixture D – 96h by MALDI-MS (3026.2 – AFPp1; 3082.2 – formyl-AFPp1). **Figure S4.** Fragmentation spectrum of formyl-PLEHHHH. **Table S1.** Formic acid hydrolysis products by HPLC-ESI-MS.

**Additional file 2.** Supplementary information 2. Reagents.

### Acknowledgements

Not applicable.

### Authors' contributions

MDM, AZ, and NG contributed to this work equally. MDM, AZ, NG, MS, AT, MRM, MF, and TK performed experiments. NG, AP, EN, and NY designed the study. MDM, EN, and NY analyzed the data and wrote the manuscript. All authors have read and approved the final manuscript.

### Funding

This research was supported by the Russian Foundation for Basic Research (project 18-29-09022\20).

### Availability of data and materials

All the data generated and/or analyzed during this study is included in this published article.

### Declarations

### Ethics approval and consent to participate

Not applicable.

### Consent for publication

Not applicable.

### Competing interests

The authors declare that they have no competing interests.

### Author details

<sup>1</sup>Biotechnology and Industrial Pharmacy Department, Lomonosov Institute of Fine Chemical Technologies, MIREA – Russian Technological University, 86 Vernadsky avenue, Moscow 119454, Russia. <sup>2</sup>Dmitry Rogachev National Medical Research Center of Pediatric Hematology, Oncology and Immunology, Laboratory of Molecular Immunology, 1 Samory Mashela street, Moscow 117997, Russia. <sup>3</sup>JSC Russian Research Center for Molecular Diagnostics and Therapy, 8 Simferopolsky boulevard, Moscow 117638, Russia. <sup>4</sup>Department of Biochemistry, Biological Faculty, Lomonosov Moscow State University, 1-12 Leninskie Gory, Moscow 119991, Russia. <sup>5</sup>I.M. Sechenov First Moscow State Medical University, 8-2 Trubetskaya street, Moscow 119991, Russia. <sup>6</sup>N. M. Emanuel Institute of Biochemical Physics, RAS. 4 Kosygina street, Moscow 119334, Russia. <sup>7</sup>Fundamentals of Biotechnology Federal Research Center, RAS, 33 Leninsky avenue, Moscow 119071, Russia. <sup>8</sup>National Research Center "Kurchatov Institute", Research Institute for Genetics and Selection of Industrial Microorganisms, 1 1-Y Dorozhnyy Proyezd, Moscow 117545, Russia.

Received: 30 July 2021 Accepted: 5 October 2021

Published online: 14 October 2021

### References

- Shen SH (1984) Multiple joined genes prevent product degradation in *Escherichia coli*. *Proc Natl Acad Sci U S A* 81:4627–4631. <https://doi.org/10.1073/PNAS.81.15.4627>
- Rozkov A, Enfors SO (2004) Analysis and control of proteolysis of recombinant proteins in *Escherichia coli*. *Adv Biochem Eng Biotechnol* 89:163–195. <https://doi.org/10.1007/b95567>
- Lee JH, Minn I, Park CB, Kim SC (1998) Acidic peptide-mediated expression of the antimicrobial peptide buforin II as tandem repeats in *Escherichia coli*. *Protein Expr Purif* 12:53–60. <https://doi.org/10.1006/PREP.1997.0814>
- Jeong DW, Shin DS, Ahn CW, Song IS, Lee HJ (2007) Expression of anti-hypertensive peptide, His-His-Leu, as tandem repeats in *Escherichia coli*. *J Microbiol Biotechnol* 17:952–959
- Hofer B (1987) Construction and stability of a sixfold repeated artificial gene. *Eur J Biochem* 167:307–313. <https://doi.org/10.1111/J.1432-1033.1987.TB13337.X>
- Li Y (2011) Recombinant production of antimicrobial peptides in *Escherichia coli*: a review. *Protein Expr Purif* 80:260–267. <https://doi.org/10.1016/j.pep.2011.08.001>
- Jenny RJ, Mann KG, Lundblad RL (2003) A critical review of the methods for cleavage of fusion proteins with thrombin and factor Xa. *Protein Expr Purif* 31:1–11. [https://doi.org/10.1016/S1046-5928\(03\)00168-2](https://doi.org/10.1016/S1046-5928(03)00168-2)
- Zorko M, Jerala R (2010) Production of recombinant antimicrobial peptides in bacteria. *Methods Mol Biol* 618:61–76. [https://doi.org/10.1007/978-1-60761-594-1\\_5](https://doi.org/10.1007/978-1-60761-594-1_5)
- Zhang J, Tian M, Chen X, Lv P, Luo W, Wang Z et al (2020) Co-expression of a thermally stable and methanol-resistant lipase and its chaperone from *Burkholderia cepacia* G63 in *Escherichia coli*. *Appl Biochem Biotechnol*:1–13. <https://doi.org/10.1007/s12010-020-03453-0>
- Jamaledine M, Harris MS, Liyanage L, Cook GA (2020) Expression, purification, and structural analysis of the full-length human integral membrane protein  $\gamma$ -sarcoglycan. *Protein Expr Purif* 167:105525. <https://doi.org/10.1016/j.pep.2019.105525>
- Zinecker S, Jakob M, Klösgen RB (2020) Functional reconstitution of TatB into the thylakoidal Tat translocase. *Biochim Biophys Acta Mol Cell Res* 1867:118606. <https://doi.org/10.1016/j.bbamcr.2019.118606>
- Partridge SM, Davis HF (1950) Preferential release of aspartic acid during the hydrolysis of proteins. *Nat* 165(4185):62–63. <https://doi.org/10.1038/165062a0>
- Hinterholzer A, Stanojlovic V, Regl C, Huber CG, Cabrele C, Schubert M (2021) Detecting aspartate isomerization and backbone cleavage after aspartate in intact proteins by NMR spectroscopy. *J Biomol NMR* 1:3. <https://doi.org/10.1007/s10858-020-00356-4>
- Pane K, Verrillo M, Avitabile A, Pizzo E, Varcamonti M, Zanfardino A et al (2018) Chemical cleavage of an Asp-Cys sequence allows efficient production of recombinant peptides with an N-terminal cysteine residue. *Bioconjug Chem* 29:1373–1383. <https://doi.org/10.1021/acs.biocconjchem.8b00083>
- Witkop B (1968) Chemical cleavage of proteins. *Science* (80-) 162:318–326. <https://doi.org/10.1126/SCIENCE.162.3851.318>
- Mizejewski GJ (2019) Protein binding and interactions with alpha-fetoprotein (AFP): a review of multiple AFP cell surface receptors, intracytoplasmic binding, and inter-molecular complexing proteins, vol 2
- Bialecki ES, Di Bisceglie AM (2005) Diagnosis of hepatocellular carcinoma. *HPB* 7:26–34. <https://doi.org/10.1080/13651820410024049>
- Tcherkassova J, Tsurkan S, Smirnova G, Borisova J, Moro R, Treshalina H (2017) Binding characterization of the targeting drug AIMPILA to AFP receptors in human tumor xenografts. *Tumor Biol* 39:1–11. <https://doi.org/10.1177/1010428317734815>
- Feldman NB, Kiselev SM, Gukasova NV, Posyanova GA, Lutsenko SV, Severin ES (2000) Antitumor activity of alpha-fetoprotein conjugate with doxorubicin in vitro and in vivo. *Biochem* 65:967–971
- Mollaev M, Gorokhovets N, Nikolskaya E, Faustova M, Zabolotsky A, Sokol M et al (2018) Recombinant alpha-fetoprotein receptor-binding domain co-expression with polyglutamate tags facilitates in vivo folding in *E. coli*. *Protein Expr Purif*. <https://doi.org/10.1016/j.pep.2017.11.001>
- Mollaev M, Gorokhovets N, Nikolskaya E, Faustova M, Zabolotsky A, Zhunina O et al (2019) Type of pH sensitive linker reveals different time-dependent intracellular localization, in vitro and in vivo efficiency in alpha-fetoprotein receptor targeted doxorubicin conjugate. *Int J Pharm*. <https://doi.org/10.1016/j.ijpharm.2018.12.073>
- Mizejewski GJ (2016) Does alpha-fetoprotein contribute to the mortality and morbidity of human hepatocellular carcinoma? A commentary. *J Hepatocell Carcinoma* 3:37–40. <https://doi.org/10.2147/jhc.s114198>

23. Zheng Y, Zhu M, Li M (2020) Effects of alpha-fetoprotein on the occurrence and progression of hepatocellular carcinoma. *J Cancer Res Clin Oncol* 146:2439–2446. <https://doi.org/10.1007/s00432-020-03331-6>
24. Moro R (2016) The alpha-fetoprotein receptor (RECAF): characterization and potential uses for cancer diagnosis and therapy. In: Lakhi N, Moretti M (eds) *Alpha-Fetoprotein Funct. Clin. Appl.* Nova Science Publishers, Inc, pp 241–275
25. Sambrook DRJ (2001) *Molecular cloning: a laboratory manual*, 4th edn. Cold spring harbor laboratory press, NY
26. Posypanova GA, Gorokhovets NV, Makarov VA, Savvateeva LV, Kireeva NN, Severin SE et al (2008) Recombinant alpha-fetoprotein C-terminal fragment: the new recombinant vector for targeted delivery. *J Drug Target* 16:321–328. <https://doi.org/10.1080/10611860801927721>
27. Landon M (1977) [15] Cleavage at aspartyl—prolyl bonds. *Methods Enzymol* 47:145–149. [https://doi.org/10.1016/0076-6879\(77\)47017-4](https://doi.org/10.1016/0076-6879(77)47017-4)
28. Ricardo J, Moro RHS (2010) Peptides that bind the alpha-fetoprotein (AFP) receptor and uses thereof. US20120270238A1
29. Bi X, Liu LF (1994) recA-independent and recA-dependent intramolecular plasmid recombination: differential homology requirement and distance effect. *J Mol Biol* 235:414–423. <https://doi.org/10.1006/JMBI.1994.1002>
30. Garnett MC (2001) Targeted drug conjugates: principles and progress. *Adv Drug Deliv Rev* 53:171–216. [https://doi.org/10.1016/S0169-409X\(01\)00227-7](https://doi.org/10.1016/S0169-409X(01)00227-7)

### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Submit your manuscript to a SpringerOpen<sup>®</sup> journal and benefit from:**

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

---

Submit your next manuscript at ▶ [springeropen.com](https://www.springeropen.com)

---