link.springer.com link.springer.com

Published online 22 September 2023 | https://doi.org/10.1007/s40843-023-2571-6
Sci China Mater 2023, 66(10): 4109–4120

Site-directed cysteine coupling of disulfide-containing non-antibody carrier proteins (THIOCAPs)

Ariana Rueda[1,2,3], Julian I. Mendoza[1,2,3], Lorena Alba-Castellon[1,2,3], Eloi Parladé[3,4,5], Eric Voltà-Durán[3,4,5], David Paez[6,7], Anna Aviño[3,8], Ramon Eritja[3,8], Esther Vázquez[3,4,5*], Antonio Villaverde[3,4,5], Ramón Mangues[1,2,3*] and Ugutz Unzueta[1,2,3,5*]

ABSTRACT **The development of a new generation of non-antibody protein drug delivery systems requires site-directed conjugation strategies to produce homogeneous, reproducible and scalable nanomedicines. For that, the genetic addition of cysteine residues into solvent-exposed positions allows the thiol-mediated cysteine coupling of therapeutic drugs into protein-based nanocarriers. However, the high reactivity of unpaired cysteine residues usually reduces protein stability, consequently imposing the use of more methodologically demanding purification procedures. This is especially relevant for disulfide-containing nanocarriers, as previously observed in THIOMABs. Moreover, although many protein scaffolds and targeting ligands are also rich in disulfide bridges, the use of these methodologies over emerging non-antibody carrier proteins has been completely neglected. Here, we report the development of a simple and straightforward procedure for a one-step production and site-directed cysteine conjugation of disulfide-containing non-antibody thiolated carrier proteins (THIOCAPs). This method is validated in a fluorescent C-X-C chemokine receptor 4 (CXCR4)-targeted multivalent nanocarrier containing two intramolecular disulfide bridges and one reactive cysteine residue strategically placed into a solvent-exposed position (THIO-T22-GFP-H6) for drug conjugation and in a humanized alternative intended for clinical applications (T22-HSNBT-H6). Thus, we produce very stable, homogeneous and fully functional antitumoral nanoconjugates (THIO-T22-GFP-H6-MMAE and T22-HSNBT-H6-MMAE) that selectively eliminate target cancer cells *via* CXCR4-receptor. Altogether, the developed methodology appears as a powerful tool for the rational engineering of emerging non-antibody, cell-targeted protein nanocarriers that contain disulfide bridges together with a solvent-exposed reactive cysteine (THIOCAP). This should pave the way for the development of a new generation of stable, homogeneous and efficient nanomedicines.**

Keywords: nanomedicine, protein carriers, disulfide-bonds, THIOCAP, nanoconjugates, cysteine coupling

INTRODUCTION

Fine-tuning of the specific drug and payload binding location is one of the main challenges in the development of a new generation of protein-based targeted drug delivery systems [1–4]. This is important since precisely controlled production procedures allow the obtention of homogeneous, reproducible, predictable and efficient nanomedical tools [5]. Therefore, site-directed drug conjugation has emerged as a promising alternative to the use of random lysine-amine-mediated binding methodologies as the ones used in the first generation of Food and Drug Administration (FDA)-approved targeted protein-based nanomedicines such as the antibody drug conjugate (ADC), Kadcyla [5–7]. For that, engineered cysteines have appeared as an interesting approach to generate coupling targets. Being this amino acid highly reactive, native cysteines are present in relatively low abundance and are usually buried or in form of disulfide bridges within non-catalytic proteins, contributing to their proper folding and thermodynamic stability [8,9]. Thus, the genetic addition of a recombinant cysteine at the solvent-exposed position of a protein makes it suitable for site-specific thiol-mediated coupling [10–12]. However, the introduction of unpaired cysteines in such exposed positions usually generates protein instability by disulfide cross-linking, resulting not only in cysteine capping but also in protein precipitation [8,13]. Although the addition of thiol reducing agents (such as dithiothreitol (DTT) or tris (2-carboxyethyl)phosphine (TCEP)) to the protein formulation could solve this problem, disulfide-containing nanocarriers such as ADCs and other non-antibody protein carriers face an additional challenge: the presence of reducing agents immediately breaks their native disulfide bridges. Importantly, such links are essential for both structural stability and functional integrity. This was already reported in the development of cysteine-coupled ADCs, known as THIOMABs [12,14,15]. Here, unpaired cysteine residues of cys-mutant antibodies were found to form disulfide interactions with other cysteine or glutathione residues. This fact caps the engineered cysteine and prevents its conjugation [12,15], while

[1] Institut d'Investigació Biomèdica Sant Pau (IIB Sant Pau), Sant Quintí 77-79, 08041 Barcelona, Spain
[2] Josep Carreras Leukaemia Research Institute (IJC Campus Sant Pau), 08041 Barcelona, Spain
[3] CIBER de Bioingeniería, Biomateriales y Nanomedicina, Instituto de Salud Carlos III, 28029 Madrid, Spain
[4] Institut de Biotecnologia i de Biomedicina, Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain
[5] Departament de Genètica i de Microbiologia, Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain
[6] Department of Medical Oncology, Hospital de la Santa Creu i Sant Pau. Barcelona, 08025 Barcelona, Spain
[7] CIBER de Enfermedades Raras (CIBERER), Instituto de Salud Carlos III, 28029 Madrid, Spain
[8] Institute for Advanced Chemistry of Catalonia (IQAC), CSIC, 08034 Barcelona, Spain
[*] Corresponding authors (emails: uunzueta@santpau.cat (Unzueta U); rmangues@santpau.cat (Mangues R); esther.vazquez@uab.es (Vázquez E))

also producing antibody instability and aggregation [16]. This issue was solved by using methodologically demanding partial reduction (using TCEP or DTT) plus re-oxidation (using $CuSO_4$ or dehydro-ascorbic acid) procedures that reduces interchain S–S bridges (as well as the engineered cysteine) to allow then, the formation of native antibody disulfide bonds [12,14–17]. However, although these procedures have allowed some cysteine-coupled ADCs to go into the clinics [18], potential undesired effects such as disulfide shuffling (also called disulfide scrambling) need to be avoided [17,19]. Moreover, these processes introduce additional steps to their industrial scale up which, following the KISS (keep it simple, stupid) principle, are meant to be kept as simple as possible [20].

Apart from that, even though drug targeting is currently mainly explored through ADCs, with around 14 products already in the market since 2000 [18,21], they still do not reach the desired clinical performance as some unsolved challenges remain. This is because although showing some targeting capacity, still a low percentage of the injected material is localized into the target cells, showing low penetrability and dose-limiting off-target toxicity [18,21–24]. Therefore, there is an increasing interest for the development of alternative non-antibody, protein-based nanomedical tools not only to allow better biodistribution of therapeutic molecules but also to achieve better penetration into target cells. In the last decades, several emerging non-antibody carrier proteins have been developed, based on both exogenous and human proteins [25–31] and have been successfully used in the first generation of targeted drug delivery systems [32–36]. However, as many of such scaffold proteins and targeting ligands contain disulfide bridges [26,31,37–40], it is mandatory to find appropriate methodological tools to allow site-directed cysteine coupling while preserving their structural and functional integrity. Since the multi-step methodologies previously described for THIOMABs have been very poorly explored in these types of emerging non-antibody nanomaterials and might also show some limitations, we aimed at developing a simpler and straightforward procedure for the one-step production and site-directed cysteine coupling of disulfide-containing non-antibody thio-carrier proteins (THIOCAPs), avoiding more demanding partial reduction and re-oxidation processes.

This has been explored in a self-assembling, protein-only multivalent nanocarrier (T22-GFP-H6) that targets the cell surface C-X-C chemokine receptor 4 (CXCR4) [40,41], which is a clinically relevant marker for tumor cells [30,42,43]. This fluorescent nanocarrier contains two intramolecular disulfide bridges that are essential for the structural integrity of T22, the targeting ligand of CXCR4 [44], and has been previously used for the targeted delivery of different antitumoral payloads into $CXCR4^+$ tumor cells in form of a first generation of random lysine-amine coupled heterogeneous nanoconjugates [33,34,36,45].

The simple and particular methodology developed here allowed the successful conjugation of a maleimide-functionalized antitumoral drug, namely monomethyl auristatin E (MMAE), to a reactive cysteine residue carefully placed into a solvent-exposed loop of the carrier protein (THIO-T22-GFP-H6). The presence of reducing agents during the purification process produced completely dysfunctional nanoconjugates with broken intramolecular disulfide bridges and a lack of targeting. However, limiting the presence of the reducing agent (TCEP) to very particular steps of the process and the following fine-tuning of the reaction stoichiometry allowed the efficient one-step purification and selective cysteine conjugation of a structurally stable and fully functional THIOCAP (THIO-T22-GFP-H6) with MMAE. The optimized method was also validated in an alternative human-derived THIOCAP designed for clinical applications (T22-HSNBT-H6) [25,32], allowing its efficient one-step purification and selective conjugation of MMAE into a solvent-exposed reactive cysteine. Optimized THIO-T22-GFP-H6-MMAE and T22-HSNBT-H6-MMAE nanoconjugates proved to be therapeutically active and highly selective for CXCR4-overexpressing ($CXCR4^+$) cancer cells.

EXPERIMENTAL SECTION

Protein design and 3D structure prediction

T22-GFP-H6, THIO-T22-GFP-H6 and T22-HSNBT-H6 proteins were designed in house and their three-dimensional (3D) structures predicted in silico by AlphaFold2 [46] algorithm integrated in ColabFold [47] from their primary FASTA sequence. Relative solvent-exposure of the native and engineered cysteines was calculated in Chimera v1.16 [48] by normalizing the surface area of the residues in the protein by the area of the same cysteines in a reference state (Gly-Cys-Gly) [49]. Main physicochemical properties of the proteins were also calculated from their primary FASTA sequences by ProtParam protein analysis Tool (Expasy) [50].

Protein production

Genes encoding for all proteins were provided by Geneart (ThermoFisher) cloned in a pET22b plasmid (Novagen) and transformed into an *Escherichia coli* (*E. coli*) Origami B (DE3, $OmpT^-$, Lon^- , $TrxB^-$, Gor^-; Novagen) by heat shock (42°C, 45 s). Cells were incubated at 37°C in presence of 100 μg mL^{-1} ampicillin, 12.5 μg mL^{-1} tetracycline and 15 μg mL^{-1} kanamycin and the temperature changed at the beginning of the exponential phase (OD_{550}: 0.5–0.7). Proteins were then produced overnight at 20°C upon induction with 0.1 mmol L^{-1} isopropyl *β-D*-1-thiogalactopyranoside (IPTG) and finally, cells were harvested by centrifugation at 5000 ×g. The percentage (%) of proteins in the cell soluble and insoluble fractions was determined by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and Western-blot immunodetection with an anti-his monoclonal antibody (Santa Cruz Biotechnology, sc-58598) upon cell disruption (by sonication) and cell fraction separation by centrifugation at 15,000 ×g.

Protein purification

Standard protocol (SP)

Cell pellets, from recombinant protein-producing bacteria, were first resuspended in aqueous buffer (20 mmol L^{-1} Tris-HCl, 500 mmol L^{-1} NaCl, 10 mmol L^{-1} imidazole, pH 8) supplemented with protease inhibitors (cOmplete EDTA free, Roche Diagnostics) and subsequently disrupted in an EmulsiFlex C5 cell disruptor (Avestin Biopharma) by two sequential rounds at 8000–10,000 psi (1 psi = 0.155 cm^{-2}). The soluble fraction of the cells was then separated by centrifugation (45 min at 20,000 ×g) and recombinant protein purified by immobilized metal affinity chromatography (IMAC) with a 1-mL HisTrap HP column (Cytiva) in an Äkta pure chromatography system (Cytiva).

Protein elution was achieved by a linear gradient of elution buffer (20 mmol L^{-1} Tris-HCl, 500 mmol L^{-1} NaCl, 500 mmol L^{-1} imidazole, pH 8), and the obtained protein was finally dialyzed against a sodium bicarbonate solution with salt (166 mmol L^{-1} $NaHCO_3$, 333 mmol L^{-1} NaCl, pH 8).

TCEP all buffer (AB) protocol

Cell pellets were resuspended in a TCEP-supplemented wash buffer (20 mmol L^{-1} Tris-HCl, 500 mmol L^{-1} NaCl, 10 mmol L^{-1} imidazole, 2 mmol L^{-1} TCEP, pH 8) in presence of protease inhibitors (cOmplete EDTA free, Roche Diagnostics) and subsequently disrupted in an EmulsiFlex C5 cell disruptor (Avestin Biopharma) by two sequential rounds at 8000–10,000 psi. The soluble fraction of the cells was then separated by centrifugation (45 min at 20,000 ×g) and subsequently charged in a 1-mL HisTrap HP IMAC column (Cytiva) with an Äkta pure chromatography system (Cytiva). The column-bound protein was then first washed with a TCEP-supplemented wash buffer (20 mmol L^{-1} Tris-HCl, 500 mmol L^{-1} NaCl, 2 mmol L^{-1} TCEP, pH 8) and elution achieved by a linear gradient of TCEP-supplemented elution buffer (20 mmol L^{-1} Tris-HCl, 500 mmol L^{-1} NaCl, 500 mmol L^{-1} imidazole, 2 mmol L^{-1} TCEP, pH 8). The obtained protein was finally thoroughly dialyzed against a sodium bicarbonate solution with salt (166 mmol L^{-1} $NaHCO_3$, 333 mmol L^{-1} NaCl, pH 8) for THIO-T22-GFP-H6 and sodium bicarbonate solution (166 mmol L^{-1} $NaHCO_3$, pH 8) for T22-HSNBT-H6 in order to completely remove any trace of TCEP.

TCEP lysis buffer (LB) protocol

Cell pellets were resuspended in a TCEP-supplemented wash buffer (20 mmol L^{-1} Tris-HCl, 500 mmol L^{-1} NaCl, 10 mmol L^{-1} imidazole, 2 mmol L^{-1} TCEP, pH 8) in presence of protease inhibitors (cOmplete EDTA free, Roche Diagnostics) and subsequently disrupted in an EmulsiFlex C5 cell disruptor (Avestin Biopharma) by two sequential rounds at 8000–10,000 psi. The soluble fraction of the cells was then separated by centrifugation (45 min at 20,000 ×g) and subsequently charged in a 1-mL HisTrap HP IMAC column (Cytiva) with an Äkta pure chromatography system (Cytiva). The column-bound protein was then washed with a TCEP-free wash buffer (20 mmol L^{-1} Tris-HCl, 500 mmol L^{-1} NaCl, pH 8) in order to remove any trace of TCEP and then eluted in a linear gradient of elution buffer (20 mmol L^{-1} Tris-HCl, 500 mmol L^{-1} NaCl, 500 mmol L^{-1} imidazole, pH 8). The obtained protein was finally dialyzed against a sodium bicarbonate solution with salt (166 mmol L^{-1} $NaHCO_3$, 333 mmol L^{-1} NaCl, pH 8) for THIO-T22-GFP-H6 and sodium bicarbonate solution (166 mmol L^{-1} $NaHCO_3$, pH 8) for T22-HSNBT-H6.

In all cases, protein purity was analyzed by SDS-PAGE and Western-blot immunodetection using an anti-His mouse monoclonal antibody (Santa Cruz Biotechnology, sc-58598). Protein integrity and presence of covalent disulfide-dimer was determined by matrix-assisted laser desorption/ionization time of flight (MALDI-TOF) mass spectrometry. The amount of purified protein was finally determined by Bradford assay.

Antitumor drug conjugation

Maleimide-functionalized MMAE was purchased at MedChemExpress (HY-15575) and resuspended in anhydrous dimethyl sulfoxide (DMSO). THIO-T22-GFP-H6 and parental T22-GFP-H6 carrier proteins were incubated in presence of different protein:drug molar ratios (1:1, 1:1.5, 1:2 and 1:50) for 4 h, in aqueous solution (166 mmol L^{-1} $NaHCO_3$, 333 mmol L^{-1} NaCl, 0.8 mmol L^{-1} EDTA pH 8) at room temperature. T22-HSNBT-H6 carrier protein was incubated at different protein: drug molar ratios of 1:1, 1:1.5 and 1:2. Finally, excess of non-conjugated MMAE molecules were completely removed from the nanoconjugates by IMAC re-purification using a HisTrap HP 1 mL column in an ÄKTA pure (Cytiva) chromatography system and dialyzed against sodium bicarbonate with salt solution (166 mmol L^{-1} $NaHCO_3$, 333 mmol L^{-1}, pH 8).

Morphometric characterization

Size distributions of THIO-T22-GFP-H6 nanoparticles and THIO-T22-GFP-H6-MMAE nanoconjugates were determined by dynamic light scattering (DLS) in a Zetasizer Ultra Red (Malvern Instruments) at 633 nm and a backscattered detector (173°). Samples were measured in triplicate (n = 3) and the average intensity size values are expressed as mean ± standard error.

Fluorescence determination

Fluorescence of T22-GFP-H6 carrier protein, THIO-T22-GFP-H6 thio-carrier protein and THIO-T22-GFP-H6-MMAE nanoconjugates (LB and AB protocols) at 0.1 mg mL^{-1} was measured in a 96-well plate reader upon excitation at 485 nm and emission detection at 535 nm. Samples were measured in triplicate.

Cell culture, protein internalization and competition assay

UM-SCC-22A-CXCR4$^+$ cells, kindly provided by Dr. Gregory Oakley and further transducted with a lentiviral CXCR4-luciferase plasmid (pLenti-III-UbC-CXCR4- 2A-luc, Abm), were cultured in 12-well plates in Dulbecco's modified Eagle's medium (DMEM, Gibco) containing 10% fetal bovine serum (FBS, Gibco) at 37°C and 5% CO_2. Cells were then incubated in presence of 20 nmol L^{-1} of THIO-T22-GFP-H6-MMAE nanoconjugates (LB and AB protocols) or parental THIO-T22-GFP-H6 and T22-GFP-H6 carrier proteins for 3 h. A potent CXCR4 receptor antagonist AMD3100 (octahydrochloride hydrate, Sigma) was incubated for 1 h before sample addition for competition assays. Cells were then washed with phosphate buffer saline (PBS) and treated with "harsh" trypsin digestion (1 mg mL^{-1}, 15 min) to remove externally attached proteins. Finally, internalized fluorescence was recorded in a MACSQuant10 flow cytometer (Miltenyi Biotec) with 488 nm laser and 525/50 nm detector. All samples were analyzed in duplicate and the results are expressed in internalized mean fluorescence ± standard error.

***In vitro* cell viability assay**

UM-SCC-22A-CXCR4$^+$ cells were cultured in 96-well plates in DMEM medium (Gibco) containing 10% FBS (Gibco) at 37°C and 5% CO_2. Cells were then incubated in presence of different concentrations of THIO-T22-GFP-H6-MMAE and T22-HSNBT-H6-MMAE nanoconjugates (0–50 nmol L^{-1}) for 48 h. Cell viability was finally tested by XTT Cell Proliferation Kit II (Roche) on a multi-well spectrophotometer at 490 nm (FLUOstar Optima, BMG, Labtech). THIO-T22-GFP-H6 and T22-HSNBT-H6 carrier proteins were added as negative controls. All samples were analyzed in triplicate and data are expressed as mean of viability ± standard error.

Statistical analysis

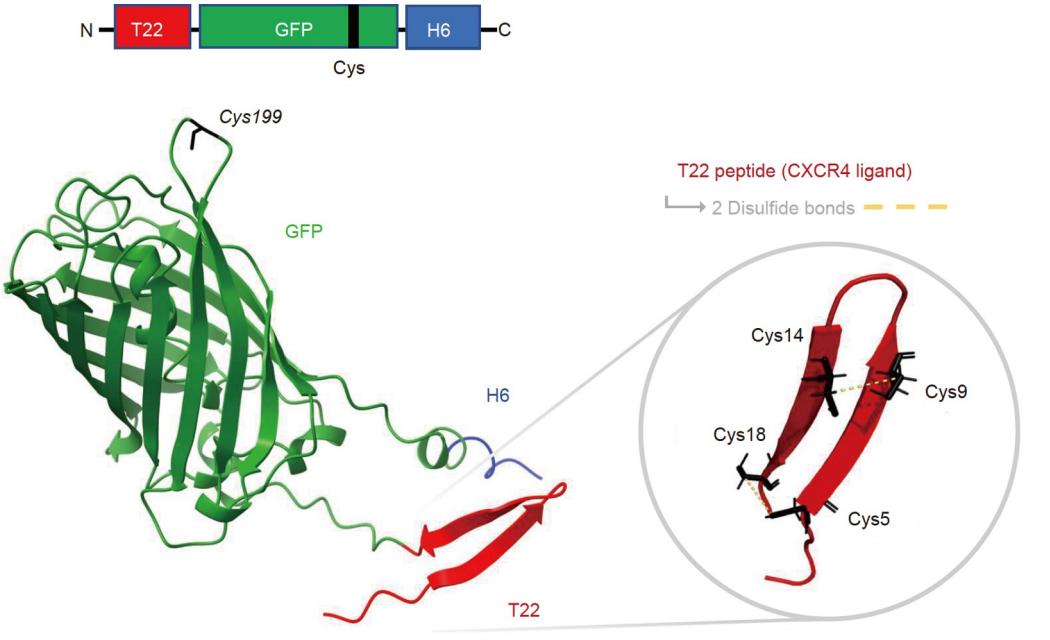
Differences in relative specific fluorescence and hydrodynamic size were assessed using one-way ANOVA with post-hoc Tukey test. Specificity studies and cell viability assays that required the comparison between two conditions were approached using multiple *t*-tests. Threshold of significance was set at $*p \leq 0.05$, $**p \leq 0.01$, and $***p \leq 0.001$. Statistical analyses were performed in GraphPad Prism 8.0.2.

RESULTS

Production and standard purification of T22-GFP-H6 based THIOCAP

T22-GFP-H6 is a 30.6-kDa disulfide-containing carrier protein that selectively targets the CXCR4 cell-surface chemokine receptor and efficiently internalizes into $CXCR4^+$ target cells *via* receptor-mediated endocytosis [40]. This fluorescent protein contains two completely buried cysteine residues within its beta-barrel structure and four solvent-exposed cysteines at the T22 ligand, which are capped in form of structural disulfide bridges (Fig. S1). Therefore, to allow the cysteine-directed site-selective conjugation of T22-GFP-H6, an additional unpaired cysteine residue was engineered at the position 199 of the carrier protein and named as THIO-T22-GFP-H6 (Fig. 1). This position was selected as it is located in a solvent-exposed loop (Fig. S1) that admits peptide insertions without affecting protein folding [51–53], and that is placed at the opposite position from T22 within the tertiary structure (Fig. 1).

Recombinant production of the novel THIOCAP (THIO-T22-GFP-H6) in *E. coli* resulted in a fully fluorescent protein that was mainly present in the soluble fraction of the cell lysate, in an amount comparable to the parental T22-GFP-H6 (Fig. 2a, Fig. S2). However, the SP of cell disruption and IMAC purification resulted in a dramatic reduction in the yield of THIO-T22-GFP-H6, which was mainly due to drastic protein precipitation induced by intermolecular disulfide cross-linking upon purification (Fig. 2a). This was reflected in the mass spectrum of the remaining soluble fraction (5%), which showed a full-length protein of 30.8 kDa and a relevant presence of covalent dimers (61.2 kDa), the combination of which had a high tendency to precipitate (Fig. 2b). In this regard, precipitated THIO-T22-GFP-H6 protein could be fully resuspended upon addition of strong reducing agents such as DTT or TCEP, resulting in stable protein monomers with low presence of dimers (Fig. 2c) but with disrupted intramolecular disulfide bridges (Fig. S3), which are essential for nanocarrier functionality.



MRRWCYRKCYKGYCYRKCRGGSSRSSSKGEE
LFTGVVPILVELDGDVNGHKFSVSGEGEGDATY
GKLTLKFICTTGKLPVPWPTLVTTLTYGVQCFS
RYPDHMKRHDFFKSAMPEGYVQERTISFKDDG
NYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILG
HKLEYNYNSHNVYITADKQKNGIKANFKIRHNIE
DCGSVQLADHYQQNTPIGDGPVLLPDNHYLST
QSALSKDPNEKRDHMVLLEFVTAAGITHGMDE
LYKHHHHHH

	T22-GFP-H6	THIO-T22-GFP-H6
NºAa	269	270
MW	30691.6	30794.7
pI	8.12	8.08
ExposedCys	4	5
S-S bridges	2	2
GRAVY	–0.710	–0.698

Figure 1 Design and physicochemical properties of THIO-T22-GFP-H6 carrier protein. Modular protein design and in silico 3D structure prediction of THIO-T22-GFP-H6 THIOCAP. Box size of the modules is only indicative. In the structure, the peptides T22 and H6 are displayed in red and blue, respectively. The inset on the right shows the intramolecular disulfide bridges present in T22. At the bottom, the amino acid sequence of the THIO-T22-GFP-H6 protein following the same color code and a table showing the main physicochemical parameters of THIO-T22-GFP-H6 and parental T22-GFP-H6 proteins are attached.

a

PROTEIN PRODUCTION	Cell soluble fraction	Cell insoluble fraction
T22-GFP-H6	94.83%	5.17%
THIO-T22-GFP-H6	93.22%	6.78%

IMAC PURIFICATION	Protein precipitation	Yield
T22-GFP-H6	3.17%	42.42 mg/L
THIO-T22-GFP-H6	94.99%	4.05 mg /L

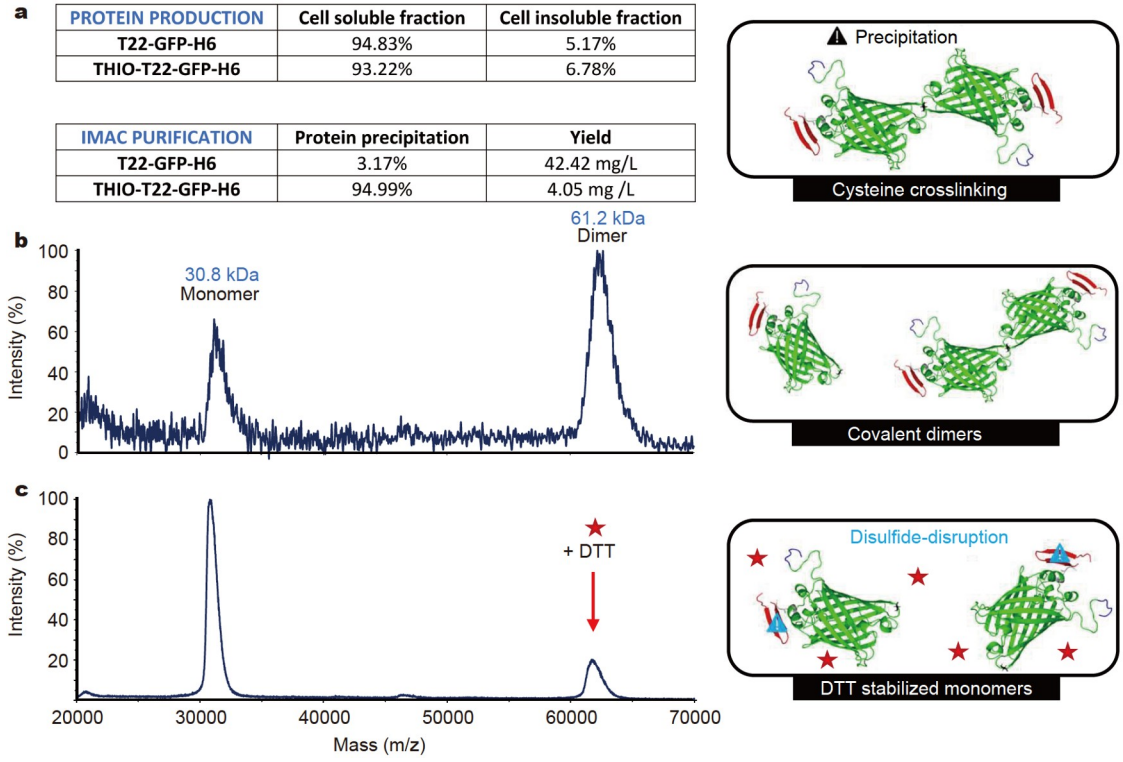


Figure 2 Recombinant production and purification of THIO-T22-GFP-H6 THIOCAP (SP). (a) Quantitative data of THIO-T22-GFP-H6 and parental T22-GFP-H6 protein recombinant production in the soluble and insoluble cell lysate fractions and percentages of protein precipitation and yield upon SP IMAC purification. Unstable disulfide dimers are depicted on the right. (b) MALDI-TOF spectrum of THIO-T22-GFP-H6 THIOCAP in the cell soluble fraction upon purification with the SP. A mixture of protein monomers and disulfide dimers are depicted on the right. (c) MALDI-TOF spectrum of THIO-T22-GFP-H6 THIOCAP upon resuspension from the cell insoluble fraction with DTT. Reducing agent-stabilized protein monomers are depicted on the right.

Fine tuning of THIO-T22-GFP-H6 purification and site-directed drug conjugation

At that point, to avoid the use of more methodologically demanding reduction and re-oxidation processes to selectively restore protein intramolecular disulfide bridges, two alternative purification protocols were explored to obtain high yields of stable and fully functional THIOCAP with intact intramolecular disulfide pairs. We first tested the addition of a reducing agent (TCEP) during the whole cell disruption and IMAC purification process, to finally remove it in the last step of the buffer exchange (AB), as previously done in other disulfide-lacking carrier proteins [10]. Then, we also tested the addition of TCEP just during a particular step of the cell lysis process to immediately remove it during the IMAC (LB) (Fig. 3a). In this context, both protocols (AB and LB) successfully stabilized THIO-T22-GFP-H6 protein by avoiding intermolecular disulfide cross-linking and subsequent precipitation, which provided protein yields comparable to that of the parental T22-GFP-H6 (Fig. 3a). Then, to evaluate if the unpaired cysteine within the THIOCAP was available for an efficient site-directed cysteine coupling, we tried to selectively conjugate a single maleimidocaproyl-functionalized MMAE, a widely used microtubule-disrupting antitumor drug [54], to THIO-T22-GFP-H6. For that, fine-tuning of the reaction stoichiometry was key to achieve an efficient but still selective Michael addition of the maleimide-drug (thioether bond) to the engineered cysteine (Cys199), while avoiding non-selective cross-reactivity (alkylamine bonds) with solvent-exposed lysine-amines as observed when using an excess of drug in high ratios (Fig. S4). In this sense, three different protein-drug molar ratios (1:1, 1:1.5 and 1:2) were tested and the obtained conjugation payloads were then analyzed by MALDI-TOF mass spectrometry (Fig. 3b). Using the LB protocol, the incubation of THIO-T22-GFP-H6 at the lowest ratios (1:1 and 1:1.5) resulted in the selective but still partial conjugation of MMAE to the Cys199 of the carrier protein (high presence of unconjugated protein). It was necessary to increase the protein-drug molar ratio up to 1:2 to observe an efficient protein conjugation with a single MMAE molecule (Fig. 3b). In contrast, no MMAE conjugation was observed in the Cys199-lacking parental T22-GFP-H6 under any of the tested conditions (Fig. 3b). Interestingly, mass spectra also revealed that the AB protocol, although being very efficient in avoiding protein precipitation, still generated THIOCAPs with broken disulfide bridges, as a heterogeneous mixture of proteins with up to 5 MMAE molecules (5 available cysteines) could be observed at all ratios (Fig. 3b). Therefore, the LB protocol revealed to be the only methodology able to produce fully folded THIO-T22-GFP-H6 THIOCAP with acceptable yields. Moreover, conjugation at a protein-drug molar ratio of 1:2 was determined to be the best stoichiometry to selectively bind a single MMAE molecule to the engineered cysteine. In this regard, an alternative maleimide-functionalized drug was also efficiently conjugated to the reactive cysteine at the optimized conditions to confirm the wide applicability of the engineered THIOCAP (Fig. S5).

Functional characterization of THIO-T22-GFP-H6-MMAE nanoconjugates

Functional characterization of the THIOCAP and generated

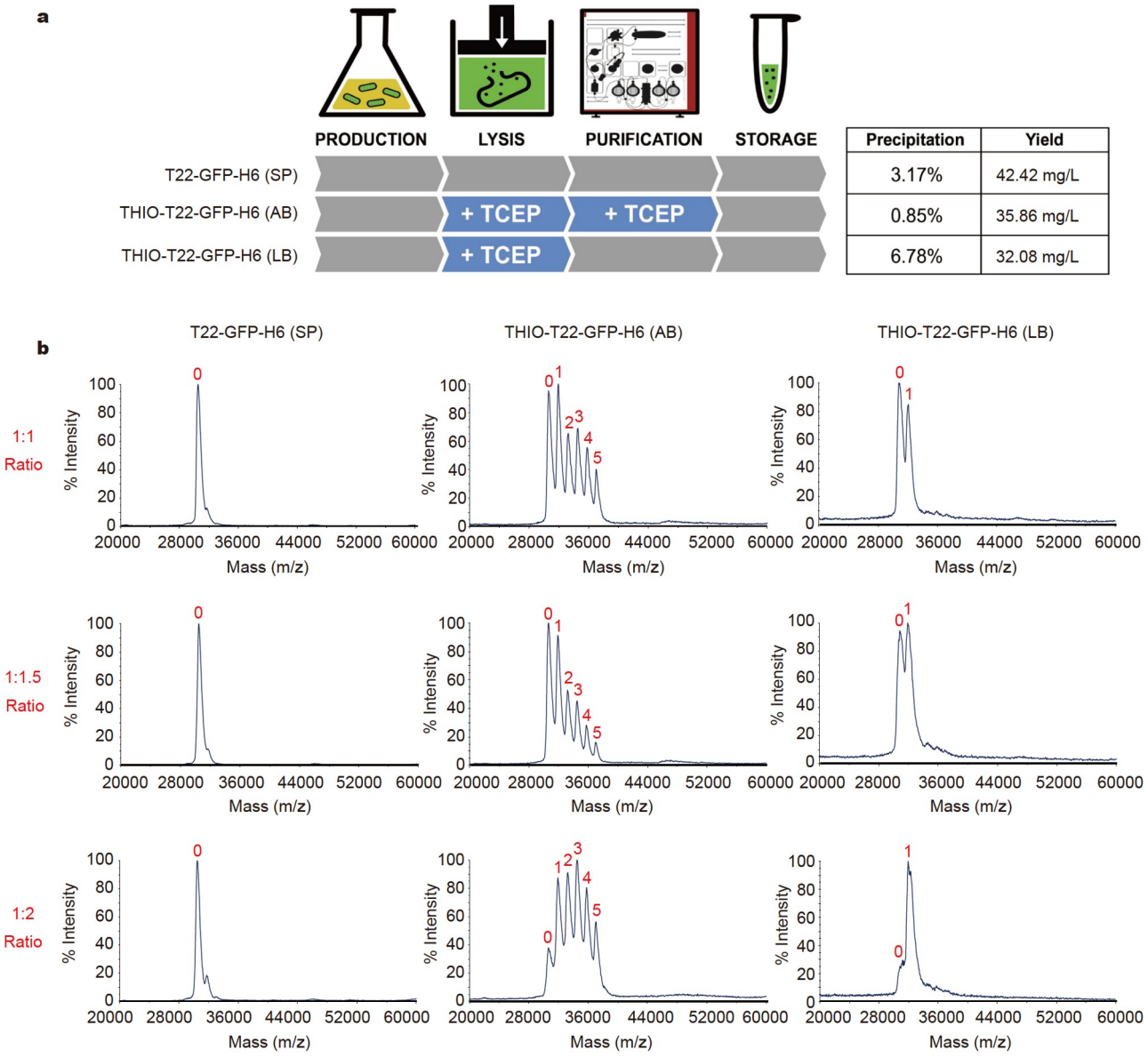


Figure 3 Fine-tuning of THIOCAP production and cysteine-directed drug conjugation. (a) Schematic representation of the THIO-T22-GFP-H6 production and purification following AB or LB protocols, in comparison with parental T22-GFP-H6 protein purified following the SP. The table shows the percentage of protein precipitation and obtained yield upon IMAC purification. (b) MALDI-TOF spectra of parental T22-GFP-H6 and THIO-T22-GFP-H6 THIOCAP produced by AB or LB method upon conjugation with different protein-drug molar ratios (1:1, 1:1.5 and 1:2) of maleimide-functionalized MMAE. Numbers above each peak indicate the amount of MMAE molecules (1.3 kDa) incorporated over the THIOCAP (30.8 kDa).

MMAE-nanoconjugates was performed in order to evaluate the suitability of the developed methodology for biomedical applications. In this sense, while the insertion of the recombinant cysteine had no impact on the tertiary structure of the carrier protein, the conjugation of MMAE into the engineered cysteine (Cys199) resulted in a mild reduction in the protein-specific fluorescence (Fig. 4a). However, generated THIO-T22-GFP-H6-MMAE nanoconjugates conserved their pseudospherical nanoscale structure with an average size distribution of around 14 nm, making them suitable for multivalent ligand presentation (Fig. 4b, Fig. S6). Of note, disulfide-disrupted MMAE-nanoconjugates (AB protocol) showed larger nanoparticle size compared with unfolded and MMAE-conjugated T22 (Fig. 4b). In consequence, THIO-T22-GFP-H6-MMAE nanoconjugates (LB protocol) efficiently internalized into CXCR4$^+$ Head and Neck UM-SCC-22A-CXCR4$^+$ cancer cells by receptor-mediated endocytosis to a similar extent as THIO-T22-GFP-H6 and parental T22-GFP-H6 nanoparticles (Fig. 4c). This fact confirmed the correct folding of intramolecular disulfide bridges within T22, the CXCR4 ligand. Related to that, internalization of the nanoconjugate was efficiently blocked by the CXCR4 antagonist AMD3100 (Fig. 4c). In contrast, disulfide-disrupted THIOCAP, generated following the AB protocol, produced completely dysfunctional nanoconjugates that failed to induce internalization into target cells (Fig. 4c). Thus, *in vitro* incubation of the functional THIO-T22-GFP-H6-MMAE nanoconjugates over UM-SCC-22A-CXCR4$^+$ cancer cells resulted in a potent and CXCR4-selective cytotoxic effect (Fig. 4d) with a half maximal

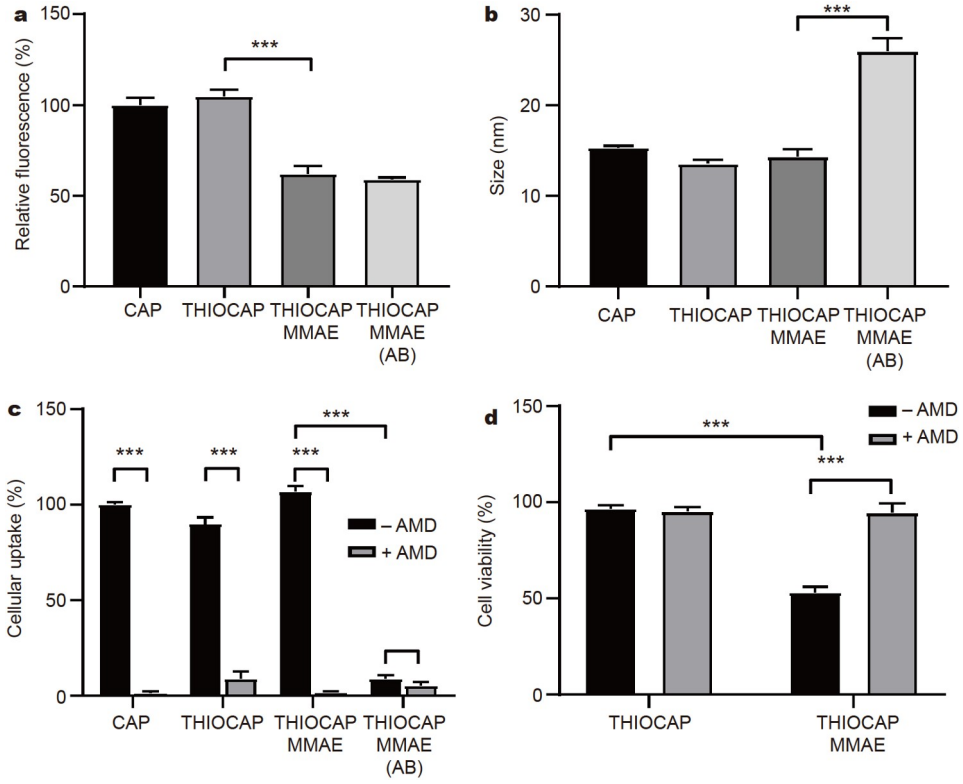


Figure 4 Functional characterization of THIO-T22-GFP-H6-MMAE nanoconjugates. (a) Relative specific fluorescence of THIO-T22-GFP-H6 thio-carrier protein (THIOCAP), functional THIO-T22-GFP-H6-MMAE nanoconjugates (THIOCAP-MMAE) and disulfide-disrupted THIO-T22-GFP-H6-MMAE nanoconjugates (THIOCAP-MMAE, AB) compared with parental T22-GFP-H6 carrier protein (CAP). (b) Size distributions of T22-GFP-H6 carrier protein (CAP), THIO-T22-GFP-H6 (THIOCAP), functional THIO-T22-GFP-H6-MMAE nanoconjugates (THIOCAP-MMAE) and disulfide-disrupted THIO-T22-GFP-H6-MMAE nanoconjugates (THIOCAP-MMAE, AB) determined by DLS. Data are represented as mean ± standard error. (c) Relative intracellular accumulation of parental T22-GFP-H6 carrier protein (CAP) and THIO-T22-GFP-H6 thio-carrier protein (THIOCAP) compared with functional THIO-T22-GFP-H6-MMAE nanoconjugates (THIOCAP-MMAE) and disulfide-disturbed THIO-T22-GFP-H6-MMAE nanoconjugates (THIOCAP-MMAE, AB) upon incubation over UM-SCC-22A-CXCR4$^+$ cancer cells at 20 nmol L^{-1} for 3 h in absence (– AMD) or presence (+ AMD) of the CXCR4 antagonist AMD3100. (d) Cytotoxic effect of the functional THIO-T22-GFP-H6-MMAE nanoconjugates (THIOCAP-MMAE) measured as UM-SCC-22A-CXCR4$^+$ cancer cell viability determined by XTT upon incubation at 20 nmol L^{-1} for 48 h in absence (– AMD) or presence (+ AMD) of the CXCR4 antagonist AMD3100. Parental THIO-T22-GFP-H6 thiolated carrier protein (THIOCAP) was used as a negative control. Significant differences between relevant data pairs are indicated as *** for $p < 0.001$.

inhibitory concentration (IC_{50}) in the nanomolar range (Fig. S7), demonstrating the full functionality and potential biomedical applications of the generated THIOCAP.

Methodology validation in a humanized THIOCAP

Lastly, the broad applicability of the developed methodology was also validated in an alternative THIOCAP derived from the human nidogen (T22-HSNBT-H6) that has been designed for biomedical applications [25,32]. T22-HSNBT-H6 is a 30.3-kDa protein that contains two structural disulfide bridges within the CXCR4 ligand T22 and an intrinsic unpaired reactive cysteine placed in a solvent-exposed loop at position 215 (Fig. S8). As expected, its recombinant production in *E. coli* following the SP resulted in a strong intermolecular disulfide cross-linking and subsequent protein precipitation upon purification (Fig. S8). Again, both protocols (AB and LB) successfully stabilized the T22-HSNBT-H6 THIOCAP by significantly reducing intermolecular disulfide cross-linking and subsequent protein precipitation (Fig. 5a). However, the MMAE conjugation again revealed that although the AB protocol is slightly more efficient in avoiding protein precipitation, it still generates disulfide-disrupted THIOCAPs, as a heterogeneous mixture of protein with up to five MMAE molecules could be observed at all the tested molar ratios (Fig. 5b). On the other hand, the LB protocol produced fully folded T22-HSNBT-H6 THIOCAPs where the unpaired cysteine (Cys 215) was fully available for its selective conjugation with a single MMAE molecule, as expected. Here, the 1:2 protein-drug molar ratio showed again to be the lowest ratio that allowed efficient site-directed cysteine conjugation while still avoiding significant unspecific lysine-amine nucleophilic attack (Fig. 5b). Finally, the generated T22-HSNBT-H6-MMAE nanoconjugates (LB protocol, 1:2 ratio) also presented a pseudospherical nanoscale structure with an average size distribution of around 32 nm (Fig. S6) and induced a strong CXCR4-dependent antitumoral effect over UM-SCC-22A-CXCR4$^+$ cancer cells (Fig. 5c) with an IC_{50} in the nanomolar range (Fig. S7).

DISCUSSION

Reactive cysteine residues fulfill a wide spectrum of functions within proteins, ranging from structural roles in disulfide folding to enzymatic activities such as redox catalysis, hydrolysis, redox signaling, metal binding or even antioxidant defense [55,56]. In consequence, unless exposed in the active site of enzymes,

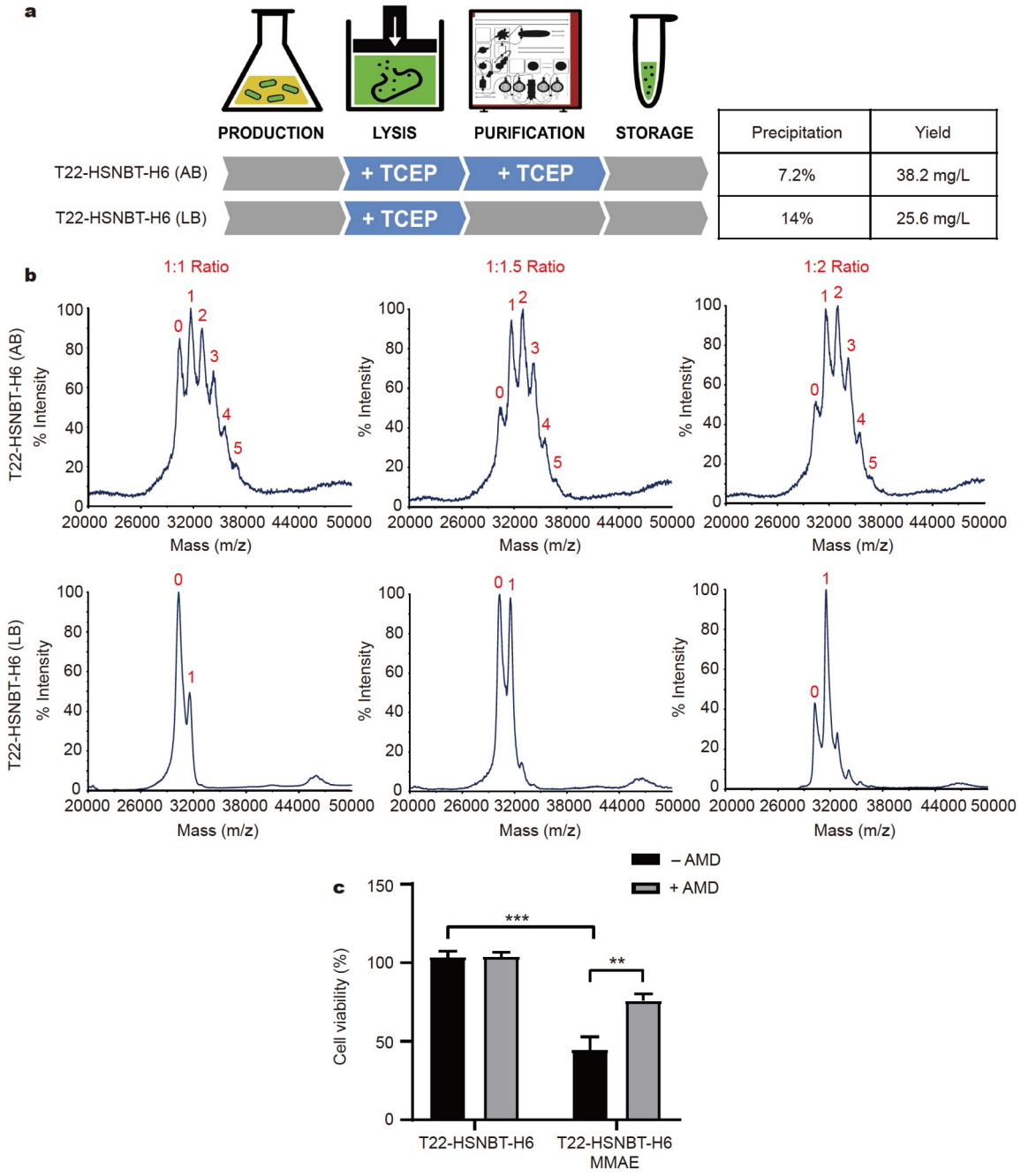


Figure 5 T22-HSNBT-H6 THIOCAP production, MMAE-conjugation and functional characterization. (a) Schematic representation of T22-HSNBT-H6 THIOCAP production and purification following the AB and LB protocols. The table shows the percentage of protein precipitation and obtained yield upon IMAC purification. (b) MALDI-TOF spectra of T22-HSNBT-H6 THIOCAP produced by AB or LB method upon conjugation with different molar ratios (1:1, 1:1.5 and 1:2) of maleimide-functionalized MMAE. Numbers above each peak indicate the amount of MMAE molecules (1.3 kDa) incorporated over the THIOCAP (30.3 kDa). (c) Cytotoxic effect of T22-HSNBT-H6-MMAE nanoconjugate (LB, 1:2 ratio) measured as UM-SCC-22A-CXCR4$^+$ cancer cell viability determined by XTT upon incubation at 20 nmol L^{-1} for 48 h in absence (– AMD) or presence (+ AMD) of the CXCR4 antagonist AMD3100. Parental T22-HSNBT-H6 THIOCAP was used as a negative control. Significant differences between relevant data pairs are indicated as *** for $p < 0.001$ and ** for $p < 0.01$.

cysteine residues are usually present in relatively low abundance and are generally capped in form of disulfide bridges or buried within the tertiary structure of proteins [57]. In this scenario, cysteines have become the primary choice to be strategically placed in a solvent-exposed position of a carrier protein to be used as a site-selective conjugation target for thiol-reacting therapeutic drugs. However, being their sulfhydryl group highly nucleophilic [57,58], introduction of unpaired cysteines into carrier proteins is usually accompanied by structural instability, aggregation and low protein yields as a consequence of inter-molecular disulfide cross-linking [10,13,16]. This was also observed in our cysteine-engineered THIOCAP variant (THIO-T22-GFP-H6), which, although being perfectly produced in the cell soluble fraction (93.22%) of *E. coli*, showed a very strong rate

of precipitation (94.99%) upon protein purification (Fig. 2a). In this sense, as both the engineered THIOCAP and the parental T22-GFP-H6 share similar physicochemical properties (Fig. 1), the observed behavior was revealed to be due to inter-molecular disulfide cross-linking (*via* introduced reactive cysteine), as generated insoluble fraction could be efficiently resuspended upon addition of a reducing agent (Fig. 2b, c).

Unlike other disulfide-lacking protein carriers that have easily solved this situation by the addition of a thiol reducing agent in the protein formulation before conjugation [59,60], disulfide-containing protein carriers such as antibodies or innovative THIOCAPs face a more challenging scenario as they need to prevent capping of engineered cysteines and intermolecular disulfide cross-linking, while maintaining intact intramolecular disulfide folding. In this sense, our novel engineered THIOCAP (THIO-T22-GFP-H6) contains one unpaired reactive cysteine strategically placed in an exposed position of the protein (Cys199) and four additional solvent-accessible cysteines, in form of two intramolecular disulfide bridges (Cys5/Cys18 and Cys9/Cys14) that are essential for its functional folding and CXCR4-targeting (Fig. 1, Fig. S1). As expected, addition of thiol-reducing agents (TCEP) to purified protein (SP) efficiently stabilized protein formulation (Fig. 2c) but also resulted in the full disulfide bridge disruption, as all five cysteines were available for MMAE conjugation (Fig. S3). Therefore, to avoid the use of more methodologically demanding partial reduction and re-oxidation procedures used in THIOMABs [12,14–17], which can potentially face some disulfide scrambling problems [17,19,61] and which have been very poorly explored in alternative disulfide-containing carrier proteins, we aimed to develop a simpler and straightforward procedure for the one-step production and site-directed cysteine-conjugation of THIOCAPs.

The addition of a thiol-reducing agent (TCEP) during the cell lysis and IMAC purification process, to be meticulously removed at the final step of the buffer exchange (AB process), drastically increased protein stability (from 94.99% to 0.85% precipitation) and the obtained yield (from 4.05 mg mL^{-1} to 35.86 mg L^{-1}) to a level comparable to that of the parental T22-GFP-H6 (Figs 2 and 3a). However, the removal of TCEP at the final step of the process, just prior to protein conjugation, was not sufficient for the efficient refolding of the native disulfides, and all five solvent-exposed cysteines were therefore available for MMAE conjugation (Fig. 3b). This generated completely dysfunctional THIO-T22-GFP-H6-MMAE nanoconjugates that failed in CXCR4-targeting and cell internalization (Fig. 4c). This was not anticipated from previous work where a non-antibody carrier protein was conjugated to an Auristatin derivative (MMAF) using a similar methodology, as they used a disulfide-lacking protein carrier [10]. On the other hand, limiting the presence of TCEP to a very particular step of the cell lysis process to be immediately removed during the IMAC purification steps (LB process) significantly improved protein stability (6.78% precipitation) and the obtained yield (32.08 mg L^{-1}), although in a slightly lesser extent than the AB process (Fig. 3a). However, such more refined methodology allowed the successful refolding of the TCEP-disrupted T22 structure to obtain a fully folded THIO-T22-GFP-H6 THIOCAP with unaltered disulfide bridges, as just the engineered cysteine (Cys199) was available for MMAE conjugation (Fig. 3b).

Then, fine-tuning of the reaction stoichiometry revealed that a slight excess of maleimide-functionalized drug (1:2 molar ratio) was necessary to obtain an efficient thio-Michael addition of the drug to Cys199 (Fig. 3b), while avoiding non-selective cross-nucleophilic attack on protein lysine-amines, as it occurs in presence of a high excess (1:50 ratio) of the drug (Fig. S4) [62,63]. As expected, the parental T22-GFP-H6, which lacks exposed cysteines, only showed non-selective lysine-amine cross-reactivity at high conjugation ratios (Fig. S4) but no MMAE coupling at lower ratios (1:1, 1:1.5 and 1:2) (Fig. 3b), corroborating the sulfhydryl-selectivity of the maleimide at low conjugation ratios under physiological conditions [63].

Thus, the particular methodology here developed allowed the efficient production and successful site-specific conjugation of a disulfide-containing THIOCAP. This approach generates a fully functional nanoconjugate (THIO-T22-GFP-H6-MMAE) that delivers its therapeutic payload into $CXCR4^+$ cancer cells (Fig. 4c) to induce a very potent and CXCR4-dependent cytotoxic effect within the nanomolar range (Fig. 4d, Fig. S7). This is achieved without preventing the cation-coordinated self-assembling capacity of the nanocarrier (Fig. 4b, Fig. S6) [64], which is very convenient for multivalent ligand presentation and receptor super-selectivity, a property highly desired in clinics [65–67].

Finally, the broader applicability of the developed methodology has been validated in an alternative disulfide-containing humanized THIOCAP (T22-HSNBT-H6), which is meant to avoid nanocarrier related immune-toxicities in targeted delivery settings [25,32]. This is highly relevant as exogenous proteins such as GFP can trigger strong immune reactions that severely limit their clinical use [68–70]. Here, the LB protocol showed again to efficiently stabilize the T22-HSNBT-H6 protein (Fig. 5a) to obtain a fully folded CXCR4-targeted THIOCAP that allowed its site-directed cysteine-coupling of a maleimide-functionalized MMAE at the 1:2 molar ratio (Fig. 5b). Then, the generated T22-HSNBT-H6-MMAE nanoparticles (Fig. S6) showed again strong CXCR4-dependent antitumoral effect (Fig. 5c) within a nanomolar range comparable to its GFP-based counterpart (Fig. S7). Moreover, the obtained high product homogeneity also paves the way for its future clinical applications, which compared with the first generation of CXCR4-targeted nanoconjugates [32,33,36,45] presents a more refined product characterization and enhanced batch-to-batch reproducibility.

CONCLUSIONS

In the context of emerging nanomedicines, we have developed here a novel and highly demanded procedure for the successful production and site-directed cysteine-conjugation of innovative, cell-targeted drug delivery systems based on disulfide-containing non-antibody thio-carrier proteins (THIOCAPs). This simple yet refined methodology has allowed the unpaired cysteine stabilization and high-yield purification of different CXCR4-targeted THIOCAPs (THIO-T22-GFP-H6 and T22-HSNBT-H6), avoiding the use of methodologically demanding partial reduction and re-oxidation procedures to conserve their intramolecular disulfide bridges, which are mandatory for their functional and conformational stability. Then, together with a fine tuning of the reaction stoichiometry, we have generated highly homogeneous and fully functional cysteine-coupled nanoconjugates that not only conserved their supramolecular nanoscale structure, but also, selectively internalized into $CXCR4^+$ cancer cells to efficiently induce a potent therapeutic action. Therefore, THIOCAP technology and its associated methodologies appear

as an emerging biomedical tool for the rational engineering of new generation of disulfide-containing non-antibody targeted nanomedicines. Moreover, their enhanced homogeneity, improved batch-to-batch reproducibility and simplified industrial scale-up will pave the way towards clinical application.

Received 5 May 2023; accepted 7 August 2023; published online 22 September 2023

1 Strop P, Liu SH, Dorywalska M, *et al.* Location matters: Site of conjugation modulates stability and pharmacokinetics of antibody drug conjugates. Chem Biol, 2013, 20: 161–167
2 Sun X, Ponte JF, Yoder NC, *et al.* Effects of drug-antibody ratio on pharmacokinetics, biodistribution, efficacy, and tolerability of antibody-maytansinoid conjugates. Bioconjugate Chem, 2017, 28: 1371–1381
3 Shen BQ, Xu K, Liu L, *et al.* Conjugation site modulates the *in vivo* stability and therapeutic activity of antibody-drug conjugates. Nat Biotechnol, 2012, 30: 184–189
4 Hamblett KJ, Senter PD, Chace DF, *et al.* Effects of drug loading on the antitumor activity of a monoclonal antibody drug conjugate. Clin Cancer Res, 2004, 10: 7063–7070
5 Sochaj AM, Świderska KW, Otlewski J. Current methods for the synthesis of homogeneous antibody-drug conjugates. Biotechnol Adv, 2015, 33: 775–784
6 Matsuda Y, Mendelsohn BA. An overview of process development for antibody-drug conjugates produced by chemical conjugation technology. Expert Opin Biol Ther, 2021, 21: 963–975
7 Chen L, Wang L, Shion H, *et al.* In-depth structural characterization of Kadcyla®(ado-trastuzumab emtansine) and its biosimilar candidate. mAbs, 2016, 8: 1210–1223
8 Petersen MTN, Jonson PH, Petersen SB. Amino acid neighbours and detailed conformational analysis of cysteines in proteins. Protein Eng Des Sel, 1999, 12: 535–548
9 Bulaj G. Formation of disulfide bonds in proteins and peptides. Biotechnol Adv, 2005, 23: 87–92
10 Goldberg SD, Cardoso RMF, Lin T, *et al.* Engineering a targeted delivery platform using Centyrins. Protein Eng Des Sel, 2016, 29: 563
11 Sadowsky JD, Pillow TH, Chen J, *et al.* Development of efficient chemistry to generate site-specific disulfide-linked protein- and peptide-payload conjugates: Application to THIOMAB antibody-drug conjugates. Bioconjugate Chem, 2017, 28: 2086–2098
12 Junutula JR, Raab H, Clark S, *et al.* Site-specific conjugation of a cytotoxic drug to an antibody improves the therapeutic index. Nat Biotechnol, 2008, 26: 925–932
13 Voynov V, Chennamsetty N, Kayser V, *et al.* Design and application of antibody cysteine variants. Bioconjugate Chem, 2010, 21: 385–392
14 Jeffrey SC, Burke PJ, Lyon RP, *et al.* A potent anti-CD70 antibody-drug conjugate combining a dimeric pyrrolobenzodiazepine drug with site-specific conjugation technology. Bioconjugate Chem, 2013, 24: 1256–1263
15 Adhikari P, Zacharias N, Ohri R, *et al.* Site-specific conjugation to cys-engineered THIOMAB™ antibodies. Methods Mol Biol, 2020, 2078: 51–69
16 Ohri R, Bhakta S, Fourie-O'Donohue A, *et al.* High-throughput cysteine scanning to identify stable antibody conjugation sites for maleimide- and disulfide-based linkers. Bioconjugate Chem, 2018, 29: 473–485
17 Zhou Q, Kyazike J, Boudanova E, *et al.* Site-specific antibody conjugation to engineered double cysteine residues. Pharmaceuticals, 2021, 14: 672
18 Fu Z, Li S, Han S, *et al.* Antibody drug conjugate: The "biological missile" for targeted cancer therapy. Sig Transduct Target Ther, 2022, 7: 93
19 Liu-Shin LPY, Fung A, Malhotra A, *et al.* Evidence of disulfide bond scrambling during production of an antibody-drug conjugate. mAbs, 2018, 10: 1190–1199
20 Mcconville FX. Scale-up dos and don'ts. In: am Ende DJ (ed.). Chemical Engineering in the Pharmaceutical Industry: R&D to Manufacturing. Hoboken, NJ: John Wiley & Sons, Inc. 2010
21 Tong JTW, Harris PWR, Brimble MA, *et al.* An insight into FDA approved antibody-drug conjugates for cancer therapy. Molecules, 2021, 26: 5847
22 Teicher BA, Chari RVJ. Antibody conjugate therapeutics: Challenges and potential. Clin Cancer Res, 2011, 17: 6389–6397
23 Abdollahpour-Alitappeh M, Lotfinia M, Gharibi T, *et al.* Antibody-drug conjugates (ADCs) for cancer therapy: Strategies, challenges, and successes. J Cell Physiol, 2019, 234: 5628–5642
24 Dean AQ, Luo S, Twomey JD, *et al.* Targeting cancer with antibody-drug conjugates: Promises and challenges. mAbs, 2021, 13: 1951427
25 Cano-Garrido O, Serna N, Unzueta U, *et al.* Protein scaffolds in human clinics. Biotechnol Adv, 2022, 61: 108032
26 Binz HK, Amstutz P, Plückthun A. Engineering novel binding proteins from nonimmunoglobulin domains. Nat Biotechnol, 2005, 23: 1257–1268
27 Unzueta U, Céspedes MV, Vázquez E, *et al.* Towards protein-based viral mimetics for cancer therapies. Trends Biotechnol, 2015, 33: 253–258
28 Casanova I, Unzueta U, Arroyo-Solera I, *et al.* Protein-driven nanomedicines in oncotherapy. Curr Opin Pharmacol, 2019, 47: 1–7
29 López-Laguna H, Sánchez-García L, Serna N, *et al.* Engineering protein nanoparticles out from components of the human microbiome. Small, 2020, 16: 2001885
30 Kim J, Takeuchi H, Lam ST, *et al.* Chemokine receptor CXCR4 expression in colorectal cancer patients increases the risk for recurrence and for poor survival. J Clin Oncol, 2005, 23: 2744–2753
31 Vazquez-Lombardi R, Phan TG, Zimmermann C, *et al.* Challenges and opportunities for non-antibody scaffold drugs. Drug Discov Today, 2015, 20: 1271–1283
32 Serna N, Pallarès V, Unzueta U, *et al.* Engineering non-antibody human proteins as efficient scaffolds for selective, receptor-targeted drug delivery. J Control Release, 2022, 343: 277–287
33 Pallarès V, Unzueta U, Falgàs A, *et al.* A multivalent Ara-C-prodrug nanoconjugate achieves selective ablation of leukemic cells in an acute myeloid leukemia mouse model. Biomaterials, 2022, 280: 121258
34 Falgàs A, Pallarès V, Unzueta U, *et al.* Specific cytotoxic effect of an auristatin nanoconjugate towards CXCR4$^+$ diffuse large B-cell lymphoma cells. Int J Nanomed, 2021, Volume 16: 1869–1888
35 Álamo P, Cedano J, Conchillo-Sole O, *et al.* Rational engineering of a human GFP-like protein scaffold for humanized targeted nanomedicines. Acta BioMater, 2021, 130: 211–222
36 Céspedes MV, Unzueta U, Aviñó A, *et al.* Selective depletion of metastatic stem cells as therapy for human colorectal cancer. EMBO Mol Med, 2018, 10: e8772
37 An FF, Zhang XH. Strategies for preparing albumin-based nanoparticles for multifunctional bioimaging and drug delivery. Theranostics, 2017, 7: 3667–3689
38 Alsultan AM, Chin DY, Howard CB, *et al.* Beyond antibodies: Development of a novel protein scaffold based on human chaperonin 10. Sci Rep, 2016, 6: 37348
39 Simeon R, Chen Z. *In vitro*-engineered non-antibody protein therapeutics. Protein Cell, 2018, 9: 3–14
40 Villaverde A, Unzueta A, Céspedes A, *et al.* Intracellular CXCR4$^+$ cell targeting with T22-empowered protein-only nanoparticles. Int J Nanomed, 2012, 7: 4533
41 Céspedes MV, Unzueta U, Álamo P, *et al.* Cancer-specific uptake of a liganded protein nanocarrier targeting aggressive CXCR4$^+$ colorectal cancer models. Nanomed-Nanotechnol Biol Med, 2016, 12: 1987–1996
42 Heras SC, Martínez-Balibrea E. CXC family of chemokines as prognostic or predictive biomarkers and possible drug targets in colorectal cancer. World J Gastroenterol, 2018, 24: 4738–4749
43 Sun X, Cheng G, Hao M, *et al.* CXCL12/CXCR4/CXCR7 chemokine axis and cancer progression. Cancer Metastasis Rev, 2010, 29: 709–722
44 Liang X. CXCR4, inhibitors and mechanisms of action. Chem Biol Drug Des, 2008, 72: 97–110
45 Pallarès V, Unzueta U, Falgàs A, *et al.* An Auristatin nanoconjugate targeting CXCR4$^+$ leukemic cells blocks acute myeloid leukemia dis-

semination. J Hematol Oncol, 2020, 13: 36
46 Jumper J, Evans R, Pritzel A, *et al.* Highly accurate protein structure prediction with AlphaFold. Nature, 2021, 596: 583–589
47 Mirdita M, Schütze K, Moriwaki Y, *et al.* ColabFold: Making protein folding accessible to all. Nat Methods, 2022, 19: 679–682
48 Pettersen EF, Goddard TD, Huang CC, *et al.* UCSF chimera? A visualization system for exploratory research and analysis. J Comput Chem, 2004, 25: 1605–1612
49 Bendell CJ, Liu S, Aumentado-Armstrong T, *et al.* Transient protein-protein interface prediction: Datasets, features, algorithms, and the RAD-T predictor. BMC BioInf, 2014, 15: 82
50 Wilkins MR, Gasteiger E, Bairoch A *et al.* Protein identification and analysis tools in the ExPASy server. In: Link AJ (ed.). 2-D Proteome Analysis Protocols. New Jersey: Humana Press, 1999. 531–552
51 Paramban RI, Bugos RC, Wen Su W. Engineering green fluorescent protein as a dual functional tag. Biotechnol Bioeng, 2004, 86: 687–697
52 Kobayashi T, Morone N, Kashiyama T, *et al.* Engineering a novel multifunctional green fluorescent protein tag for a wide variety of protein research. PLoS ONE, 2008, 3: e3822
53 Voltà-Durán E, Cano-Garrido O, Serna N, *et al.* Controlling self-assembling and tumor cell-targeting of protein-only nanoparticles through modular protein engineering. Sci China Mater, 2020, 63: 147–156
54 Younes A, Yasothan U, Kirkpatrick P. Brentuximab vedotin. Nat Rev Drug Discov, 2012, 11: 19–20
55 Netto LES, de Oliveira MA, Monteiro G, *et al.* Reactive cysteine in proteins: Protein folding, antioxidant defense, redox signaling and more. Comp Biochem Physiol Part C-Toxicol Pharmacol, 2007, 146: 180–193
56 Giles NM, Giles GI, Jacob C. Multiple roles of cysteine in biocatalysis. Biochem Biophys Res Commun, 2003, 300: 1–4
57 Gunnoo SB, Madder A. Chemical protein modification through cysteine. ChemBioChem, 2016, 17: 529–553
58 Spicer CD, Davis BG. Selective chemical protein modification. Nat Commun, 2014, 5: 4740
59 Mahalingam SM, Dudkin VY, Goldberg S, *et al.* Evaluation of a centyrin-based near-infrared probe for fluorescence-guided surgery of epidermal growth factor receptor positive tumors. Bioconjugate Chem, 2017, 28: 2865–2873
60 Klein D, Goldberg S, Theile CS, *et al.* Centyrin ligands for extrahepatic delivery of siRNA. Mol Ther, 2021, 29: 2053–2066
61 Moritz B, Stracke JO. Assessment of disulfide and hinge modifications in monoclonal antibodies. Electrophoresis, 2017, 38: 769–785
62 Brewer CF, Riehm JP. Evidence for possible nonspecific reactions between *N*-ethylmaleimide and proteins. Anal Biochem, 1967, 18: 248–255
63 Ravasco JMJM, Faustino H, Trindade A, *et al.* Bioconjugation with maleimides: A useful tool for chemical biology. Chem Eur J, 2019, 25: 43–59
64 López-Laguna H, Unzueta U, Conchillo-Solé O, *et al.* Assembly of histidine-rich protein materials controlled through divalent cations. Acta BioMater, 2019, 83: 257–264
65 Martinez-Veracoechea FJ, Frenkel D. Designing super selectivity in multivalent nano-particle binding. Proc Natl Acad Sci USA, 2011, 108: 10963–10968
66 Carlson CB, Mowery P, Owen RM, *et al.* Selective tumor cell targeting using low-affinity, multivalent interactions. ACS Chem Biol, 2007, 2: 119–127
67 Angioletti-Uberti S. Theory, simulations and the design of functionalized nanoparticles for biomedical applications: A soft matter perspective. npj Comput Mater, 2017, 3: 48
68 Gambotto A, Dworacki G, Cicinnati V, *et al.* Immunogenicity of enhanced green fluorescent protein (EGFP) in BALB/c mice: Identification of an H2-K^{d}-restricted CTL epitope. Gene Ther, 2000, 7: 2036–2040
69 Ansari AM, Ahmed AK, Matsangos AE, *et al.* Cellular GFP toxicity and immunogenicity: Potential confounders in *in vivo* cell tracking experiments. Stem Cell Rev Rep, 2016, 12: 553–559
70 Morris JC, Conerly M, Thomasson B, *et al.* Induction of cytotoxic T-lymphocyte responses to enhanced green and yellow fluorescent proteins after myeloablative conditioning. Blood, 2004, 103: 492–499

Acknowledgements The authors are indebted to ISCIII (PI20/00400) co-funded by European Regional Development Fund (ERDF, a way to make Europe), and to CIBER-BBN (project NANOSCAPE and NANOLINK) granted to Unzueta U; to AEI (PID2019-105416RB-I00/AEI/10.13039/501100011033) and to CIBER-BBN (NANOREMOTE) granted to Vázquez E; to ISCIII (PI21/00150) co-funded by European Regional Development Fund (ERDF, a way to make Europe), to CIBER-BBN (4NanoMets) and to AGAUR (2021 SGR-01140) granted to Mangues R; and to CIBER-BBN (VENOM4CANCER) and AGAUR (2021 SGR-00092) granted to Villaverde A. This work was also supported by CIBER-Consorcio Centro de Investigación Biomédica en Red- (CB06/01/1031 and CB06/01/0014), Instituto de Salud Carlos III, Ministerio de Ciencia e Innovación and European Regional Development Fund (ERDF) and by CERCA programme (Generalitat de Catalunya). Unzueta U was supported by Miguel Servet contract (CP19/00028) from ISCIII co-funded by European Social Fund (ESF investing in your future). Alba-Castellon L was supported by the Spanish Association Against Cancer-AECC (POSTD20070ALBA). Rueda A was supported by a PFIS predoctoral fellowship (FI21/00012) from ISCIII co-funded by European Social Fund (ESF, investing in your future) and EVD by a predoctoral fellowship from Ministerio de Ciencia, Innovación y Universidades (FPU18/04615). Villaverde A received an Icrea Academia award. Protein production was partially performed by the ICTS "NANBIOSIS", more specifically by the Protein Production Platform of CIBER in Bioengineering, Biomaterials & Nanomedicine (CIBER-BBN)/ IBB, at the UAB http://www.nanbiosis.es/portfolio/u1-protein-production-platform-ppp/. Molecular graphics were performed with UCSF Chimera and ChimeraX, developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from the National Institutes of Health R01-GM129325 and the Office of Cyber Infrastructure and Computational Biology, National Institute of Allergy and Infectious Diseases.

Author contributions Rueda A performed most of the experiments and prepared the figures; Mendoza JI prepared the humanized nanoconjugate and prepared the figures; Alba-Castellon L performed *in-vitro* cytotoxicity assays and prepared the figures; Voltà-Durán E and Parladé E performed in-silico analysis and prepared the figures. Paez D, Aviño A and Eritja R analyzed the data and supervised the chemical part of the study. Vázquez E, Mangues R, Villaverde A and Unzueta U conceived and supervised the whole study; Unzueta U prepared the first draft of the manuscript. All authors contributed to the general discussion and approved the manuscript.

Conflict of interest The authors declare that they have no conflict of interest. Mangues R, Vazquez E and Villaverde A are co-founders of Nanoligent SL, a company devoted to developing nanostructured drugs for cancer treatment. Mangues R, Vázquez E, Villaverde A, Parladé E and Unzueta U are inventors of a patent covering the use of HSNBT for biomedical applications.

Supplementary information Experimental details and supporting data are available in the online version of the paper.

Ariana Rueda graduated in biotechnology in 2020 and obtained a Master's degree in biochemistry, molecular biology and biomedicine. She is currently pursuing a PhD degree in biotechnology directed by Dr. Ugutz Unzueta and Dr. Ramón Mangues at the Sant Pau Biomedical Research Institute of Barcelona, Spain. Her research focuses on developing new protein-based drug nanoconjugates (nanomedicines) for targeted therapies in the field of oncology.

Esther Vázquez obtained an M.D. PhD degree at the University of Valladolid and completed her training at the University of Oviedo, the State University of New York and the Institute of Cancer Research in London. She is currently a Researcher at the Autonomous University of Barcelona. She leads a team focused on the development of new protein-based materials that serve as drug delivery systems, formed by multidomain proteins that self-assemble into nanoparticles for targeted therapy in cancer.

Ugutz Unzueta graduated in biotechnology in 2008 and got his PhD degree in 2013. His research career has been developed in Spain, United Kingdom and Italy. From 2020, he has been leading an independent research line dedicated to the design of self-assembling protein materials and drug nanoconjugates for precision nanomedicines at Sant Pau Biomedical Research Institute (Barcelona, Spain). He is also a member of the Spanish nanomedicine network CIBER-BBN and associated professor at the Autonomous University of Barcelona.

Ramón Mangues became a pharmacist (1981), specialist in hospital pharmacy (1987) and PharmD PhD (1988) at Navarra University in Spain. He was a postdoctoral (1988–1992) and associated (1992–1998) researcher at New York University Medical Center. He became a National Health System Researcher at Sant Pau Biomedical Research Institute (Spain), where he leads the Group of Oncogenesis and Antitumor Drugs focusing on targeted drug delivery to selectively eliminate cancer stem cells to obtain antimetastic effect.