DE GRUYTER DOI 10.1515/hsz-2013-0116 — Biol. Chem. 2013; 394(8): 995–1004

Review

Hauke Lilie*, Susanne Richter, Sabine Bergelt, Stefan Frost and Franziska Gehle

Polyionic and cysteine-containing fusion peptides as versatile protein tags

Abstract: In response to advances in proteomics research and the use of proteins in medical and biotechnological applications, recombinant protein production and the design of specific protein characteristics and functions has become a widely used technology. In this context, protein fusion tags have been developed as indispensable tools for protein expression, purification, and the design of functionalized surfaces or artificially bifunctional proteins. Here we summarize how positively or negatively charged polyionic fusion peptides with or without an additional cysteine can be used as protein tags for protein expression and purification, for matrix-assisted refolding of aggregated protein, and for coupling of proteins either to technologically relevant matrices or to other proteins. In this context we used cysteine-containing polyionic fusion peptides for the design of immunotoxins. In general, polyionic fusion tags can be considered as a multifunctional module in protein technology.

Keywords: bifunctional proteins; disulfide bonds; immunotoxins; polyionic interactions; protein purification; renaturation.

***Corresponding author: Hauke Lilie,** Institut für Biochemie und Biotechnologie, Martin-Luther-Universität Halle-Wittenberg, Kurt-Mothes Strasse 3, D-06120 Halle/Saale, Germany, e-mail: hauke.lilie@biochemtech.uni-halle.de
Susanne Richter: DSP Development Cell Culture, Sandoz GmbH, Sandoz Biopharmaceuticals Schaftenau, Biochemiestrasse 10, A-6336 Langkampfen, Austria
Sabine Bergelt: Gene Center, Ludwig-Maximilians-Universität (LMU) Munich, Department of Biochemistry, Feodor-Lynen-Strasse 25, 81377 Munich, Germany
Stefan Frost: Department of Chemistry, Umeå University, SE-901 87 Umeå, Sweden
Franziska Gehle: Institut für Biochemie und Biotechnologie, Martin-Luther-Universität Halle-Wittenberg, Kurt-Mothes Strasse 3, D-06120 Halle/Saale, Germany

Introduction

The ever increasing demand on recombinant protein production and purification, as well as the application of proteins in different pharmacological and technological set-ups, inspired the development of a large number of protein tags that can be used in protein production and/or application. These protein tags range in molecular size and complexity from small peptides of only a few amino acids to large proteins such as thioredoxin, maltose-binding protein (MBP) or glutathione-S-transferase (GST). Usually these large protein fusions facilitate higher expression yields of the target protein, enhance protein stability against proteolysis, improve solubility, and enable an affinity purification of the fusion protein. A comprehensive summary of the characteristics and the use of such fusion proteins can be found in a number of recent reviews (Hearn and Acosta, 2001; Terpe, 2003; Young et al., 2012). Even though these large fusion partners are well established in protein chemistry, they may possess a disadvantageous effect on the function or oligomerization of the target protein. Thus, these fusion partners are almost always removed by proteolysis after purification.

Small peptide tags interfere to a lesser extent with protein structure and function. Indeed, approximately 60% of all protein structures deposited in the pdb databank contain a His-tag (Derewenda, 2004). The His_6-tag and Strep tag II (Schmidt and Skerra, 2007) are probably the peptidic protein tags most often used in recombinant protein purification. Furthermore, these tags can be used for protein detection, as there are antibodies commercially available that specifically bind these short peptides. This kind of antibody-mediated detection of fusion proteins is also used for other peptide fusion tags, such as the FLAG-tag or the c-myc-tag (Terpe, 2003; Young et al., 2012).

In addition to purification and detection, the His-tag has also been exploited for matrix-assisted refolding of proteins (Sinha et al., 1994; Negro et al., 1997; Rogl et al., 1998; Zouhar et al., 1999). Overexpression of recombinant proteins often leads to accumulation of the protein in the form of inactive inclusion bodies (Marston, 1986). In order to transfer this protein from the inactive into the native and active state, these inclusion bodies must be solubilized and refolded. A major side reaction of this refolding process *in vitro* is protein aggregation (Kiefhaber et al., 1991), which can be reduced by coupling the solubilized

protein to a matrix prior to refolding. Using this strategy, protein refolding can be performed at comparably high protein concentrations and in a technically very-well-defined process.

The primary function of most of the protein tags is the ability to specifically bind to a certain target structure. These can be Ni-NTA (His-tag), a streptavidin variant (strep tag II), GSH (GST), amylose (MBP), calmodulin (calmodulin-binding-peptide) and others (Hearn and Acosta, 2001; Young et al., 2012). This binding affinity can be used for functionalization of a matrix by immobilizing an appropriately tagged protein with the desired associated function, e.g. an enzymatic activity. A very high affinity of a fusion protein to a matrix and thus a very stable immobilization has been achieved via coiled-coil peptides (Chao et al., 1998). Based on this kind of protein technology, functionalized surfaces and nanoparticles have been generated for biosensor and immunological applications (Chao et al., 1998; Craig et al., 2012).

In this review we describe the different possible applications of polyionic fusion peptides containing a stretch of either positively charged amino acids (arginine or lysine) or negatively charged amino acids (glutamic acid or aspartic acid) in protein technology. These polyionic fusion peptides might additionally contain a cysteine residue, which can transfer non-covalent polyionic interactions into a disulfide-bonded covalent coupling reaction. Such polyionic fusion peptides have been applied in protein purification, immobilization, matrix-assisted refolding, and the covalent coupling of different proteins in order to design artificially bifunctional proteins.

Polyionic peptide tags in expression and purification of recombinant proteins

Whenever fusing a tag to a protein, the first decision to be made is whether the tag should be positioned at the N- or C-terminus of the target protein. Obviously, this decision depends on the structural accessibility and solvent exposure of both termini. A statistical analysis of structures of a small subset of monomeric proteins in the pdb databank showed that in 80.3% of the cases the N-terminus is solvent-accessible and in 86.1% the C-terminus is accessible to solvent (Jacob and Unger, 2007). In the case of domains of larger proteins, the ideal position of a tag is the terminus of the domain that would be fused to the rest of the natural full-length protein. At the genetic level, the actual coding sequence near the start codon is important for translation initiation and thus affects the expression level of recombinant proteins (Jana and Deb, 2005). During recombinant expression of yeast α-glucosidase in *E. coli*, an N-terminally positioned polyionic fusion tag decreased the production of the enzyme more than ten-fold, whereas the same polyionic tag at the C-terminus did not influence, or only slightly influenced, the expression yield compared to the non-modified enzyme (Richter and Lilie, unpublished). If the polyionic tag has to be placed N-terminally, e.g. because of functional reasons, then it might be very useful to combine it with another tag, such as a cleavable GST (Kleinschmidt et al., 2003) or a His-tagged SUMO fusion. In such a construct, the polyionic tag is placed in the middle of the fusion protein, thus not affecting expression any more. Purification can subsequently be performed very efficiently, utilizing a two-tag strategy (Kleinschmidt et al., 2003).

An alternative to positioning the polyionic tag at either the N- or the C-terminus of the protein, is the integration of the fusion peptide into a surface-exposed loop of the protein, as has been done for the viral coat protein VP1 of polyoma virus (Stubenrauch et al., 2000). This strategy, however, requires a detailed structural and functional knowledge of the protein of interest.

To date, all proteins containing a polyionic tag have been produced in the cytoplasm of *E. coli*. Therefore, there are currently no data available on the effect of a polyionic tag on the secretion of a recombinant protein into the periplasm. One might assume that a C-terminal fusion of such a highly charged sequence probably does not interfere with the interaction of the protein to be secreted with SecA and SecB, which is the initial reaction of protein transport across the cytoplasmic membrane via the Sec pathway; however, this is currently rather speculative.

Polycationic fusion peptides

Sassenfeld and Brewer (1984) were the first to use a polyarginine tag for protein purification. They fused a penta-arginine-tag to the C-terminus of β-urogastrone, over-produced the protein in *E. coli* and purified it to 95% homogeneity via two steps of ion-exchange chromatography. The first purification step was a cation exchange, and because most *E. coli* proteins are neutral or slightly acidic, they do not bind to the resin. In contrast, the polyarginine tag enables a strong interaction of the recombinant protein with the matrix. The eluted protein was then digested with carboxypeptidase B to remove the arginine-tag and the tag-free protein was subsequently purified by an anion-exchange chromatography step.

Alternatively, it would have been possible to re-chromatograph the tag-free protein on the cation exchange resin, as it would no longer bind to the column or would elute at a lower ionic strength. Thus, the possibility to remove by proteolysis of the C-terminal tag using carboxypeptidase B greatly facilitates efficient protein purification. Initially, it was proposed that the proteolytic removal of the tag is neither efficient nor very specific (Nagai and Thogersen, 1984), however, a detailed analysis of polyarginine-tagged RNase A demonstrated perfectly controlled and highly efficient removal of the tag (Fuchs and Raines, 2005).

The interaction of a polyarginine-tagged protein with an ion-exchange resin strongly depends on the amount of positive charges on the tag. Although no systematic analyses have been performed, nevertheless a comparison of published data shows that such a correlation exists. RNase A tagged with nine arginine residues eluted from a SP-sepharose at a salt concentration >1 M NaCl (Fuchs and Raines, 2005), whereas a tag of eight or six arginine residues conferred a binding to cation exchange resins at concentrations up to 450 mM and 80 mM NaCl, respectively (Stempfer et al., 1996; Stubenrauch et al., 2001). Obviously, the interaction of polyarginine-tagged proteins with a resin does not only depend on the length of the tag but also on the structure of the matrix. A polyionic matrix such as heparin-sepharose would be much more efficient in binding polyarginine peptides than conventional ion-exchange matrices.

Instead of a chromatographic purification of polyarginine-tagged proteins via ion-exchange chromatography, these protein fusions can also be purified from a crude cell extract via precipitation with the polyelectrolyte polyacrylic acid (Ford et al., 1991).

If the tag is needed after purification for any further applications (see below) then it can be protected against proteolytic activity of several carboxypeptidases by introducing an amino acid such as proline at the C-terminus, which is not a substrate for these proteases. The polyionic tag should not interfere with structure, function, and stability of the fusion protein. We have never observed any negative effect of our polyarginine tag (8–10 arginines) on the fused enzymes or antibody fragments tested (Richter et al., 2001, Kleinschmidt et al., 2003, Bergelt et al., 2009). The crystal structure of an arginine-tagged maltodextrin-binding protein was identical to that of the non-tagged protein (Bucher et al., 2002). Although the enzyme activity of a RNase A tagged with nine arginines was also unaltered, the stability of the tagged protein was slightly decreased compared to the wildtype protein because of an electrostatic repulsion of the positively charged fusion peptide with the positively charged protein surface (Fuchs and Raines, 2005). However, as for any other peptide tag, the influence of polyionic tags on stability and activity of any single fusion protein needs to be checked experimentally.

In principle, polylysine can also be exploited as a polyionic fusion tag. Indeed, for the chemically synthesized peptide NPY we used an octalysine tag for protein coupling (see below) and a fusion protein of glucagon-like peptide-1 has been refolded and purified successfully on a cation exchange resin using a hexalysine tag (Kim et al., 2012). For recombinantly produced proteins, however, polyarginine tags have usually been described. Favoring arginine over lysine in these polyionic tags is probably based on genetic reasons. Lysine is encoded by only two different codons whereas there are six codons for arginine. Even though only two arginine codons are used frequently in *E. coli*, currently available *E. coli* expression strains are supplemented with additional $tRNA^{Arg}$. This makes it easier to design the encoding sequence of a polyarginine tag.

Polyanionic fusion peptides

As an alternative to polyarginine fusion peptides, negatively charged polyglutamate/aspartate tags can be exploited. Polyglutamate/aspartate tags are less sensitive to proteolytic degradation than polyarginine (Parker et al., 1990; Zhao et al., 1990). Initially these polyanionic fusion peptides, consisting of up to eleven aspartate residues, were designed for the purification of the respective fusion proteins via polyelectrolyte precipitation using polyethyleneimine (Parker et al., 1990; Zhao et al., 1990). However, Stubenrauch et al. (2000) and Richter et al. (2001) could demonstrate that a tag of 8–10 glutamic acid residues, either introduced in a loop of the polyoma viral protein VP1 or placed at the C-terminus of the Fd chain of an antibody Fab fragment, permitted an efficient purification by anion-exchange chromatography.

Matrix-assisted refolding of polyionically tagged proteins

Renaturation of proteins *in vitro* from inclusion bodies obtained by recombinant production can be a very tedious process, as a number of biophysical factors have to be considered (De Bernardez Clark et al., 1999). The predominant side reaction limiting the renaturation yield is protein aggregation. As aggregation is a higher order reaction, it can be suppressed very efficiently by

reversible immobilization of the denatured protein on a solid support prior to renaturation. This concept was initially applied by Creighton (1985), who used an ion-exchange resin. The denatured protein was bound to a commercially available ion-exchange column in a buffer of low ionic strength containing a high concentration of urea. By decreasing the urea concentration gradually, the bound protein folded to its native state and could be eluted subsequently by increasing the ionic strength of the buffer. Although this strategy led to a successful matrix-assisted refolding of the target protein, it is limited by a conceptual problem: because of classical interaction of a protein to an ion-exchange resin the protein is bound via multiple attachment sites depending on the number of charged amino acids in the protein sequence. This multivalent binding might have a strong negative effect on refolding. To circumvent this problem, Stempfer et al. (1996a) used a hexaarginine tag fused to the N- or C-terminus of α-glucosidase to immobilize the denatured protein on heparin-sepharose. To ensure that the protein is only attached to the matrix via the polyionic tag, a salt concentration of 40 mM NaCl was added to the buffer. Higher salt concentrations caused unspecific hydrophobic interactions of the protein with the matrix, which again limited the yield of refolded protein. Whereas the arginine-tag showed a negative influence on refolding of α-glucosidase in solution compared to the non-tagged wildtype protein, the matrix-assisted refolding of the hexaarginine-tagged α-glucosidase was more efficient and, importantly, could be performed at a 100-fold higher protein concentration (Figure 1A). Subsequently, this concept of matrix-assisted refolding on an ion-exchange resin could be transferred to a protein tagged with a decaglutamic acid sequence and bound to an anion-exchange matrix (Figure 1B). These data showed that: (i) the polyglutamate tag increased the yield of refolding already with the soluble protein; (ii) similar to the arginine-tagged variant, matrix-assisted refolding could be performed at a 100-fold higher protein concentration compared to refolding in solution; and (iii) the success of refolding strongly depends on the matrix.

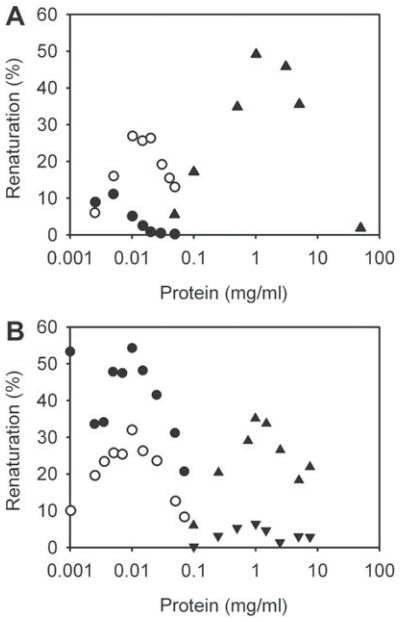

Figure 1 Matrix-assisted refolding of α-glucosidase.
(A) Yield of renaturation of wildtype α-glucosidase and α-glucosidase fused to a hexaarginine tag at its C-terminus. Renaturation was performed in 10 mM phosphate pH 7.6, 30 mM NaCl, 8% (w/v) ethylene glycol, 10°C either in solution (open circles wildtype α-glucosidase; closed circles α -glucosidase-Arg_6) or with the protein immobilized on heparin-sepharose (closed triangles). Data taken from Stempfer et al., 1996b.(B) Yield of renaturation of wildtype α-glucosidase and α-glucosidase fused to a decaglutamate tag at its C-terminus. Renaturation was performed in 10 mM phosphate pH 7.6, 1 mM EDTA, 10°C either in solution (open circles wildtype α-glucosidase; closed circles α-glucosidase-Glu_{10}) or with the protein immobilized on Fractoprep DEAE (closed triangles up) or Q-sepharose (closed triangles down). Richter and Lilie, unpublished results.

Immobilization of proteins via polyionic fusion peptides

Polyionic fusion peptides are very well suited for binding of fusion proteins to charged resins for either protein purification or matrix-assisted renaturation. Depending on the number of charged amino acids on the tag and the structure of the solid support, this interaction can be very stable; elution of RNase A containing a nona-arginine-tag from a SP-sepharose required a NaCl concentration >1 M (Fuchs and Raines, 2005). This strong interaction can be used to establish a stable non-covalent immobilization on such a matrix. In this way Stempfer et al. (1996b) designed an enzyme reactor in which they immobilized the enzyme α-glucosidase with C-terminal polycationic hexaarginine fusion peptide to heparin-sepharose. The immobilized α-glucosidase exhibited the same stability toward thermal or urea-induced denaturation, and the same enzymatic activity as the soluble enzyme. However, the operational stability of the coupled enzyme under conditions of continuous substrate conversion was much higher compared to the soluble form. In contrast to covalent immobilization, the non-covalent immobilization allowed regeneration of the enzyme reactor by simply eluting the enzyme at

high salt concentrations and subsequently re-charging the matrix with fresh enzyme.

Another interesting application of protein immobilization is the functionalization of flat surfaces. The uniform and oriented immobilization of proteins is important in the design of biosensors, supramolecular nanostructures or for formation of two-dimensional crystals used in structural analyses by electron and scanning probe microscopy. One material well established for such analyses is mica. Because of the atomic structure of its aluminosilicate crystals, mica possesses a very flat, negatively charged hydrophilic surface. By fusing either green fluorescent protein or GST to a hexaarginine tag, Nock et al. (1997) could specifically and reversibly bind these fusion proteins to mica plates. In analogy to the situation described for matrix-assisted refolding, immobilization necessitated that the ionic strength of the buffer be adjusted in order to minimize random, non-specific electrostatic interactions of the target protein with the surface and to achieve an attachment where the protein is only bound by the polyionic fusion peptide.

Design of artificial bifunctional proteins via association of polyionic fusion peptides

Proteins are indispensible tools in biotechnology, medicine, and basic research and their different functions are exploited in many kinds of applications. Modern molecular biology provides the opportunity to design and produce tailor-made proteins with functions or molecular properties not found in natural proteins. Combining functions of different proteins is of particular interest; well-established examples are enzyme-linked antibodies for ELISA diagnostics or the fusion of protein toxins to tumor-specific antibody fragments as therapeutically active immunotoxins (Pastan et al., 2007; Weldon and Pastan, 2011). Oppositely charged polyionic fusion peptides can be used as biochemical tools to create such bifunctional protein complexes from the appropriately tagged proteins.

Non-covalent heterodimer formation of polyarginine- and polyglutamate-tagged proteins

Binding of polyionically tagged proteins to ion-exchange resins has been shown to be very stable depending on the ionic strength of the buffer. For these interactions, polyionic matrices such as heparin-sepharose, or on solid supports with charged groups positioned on flexible arms such as tentaclegels, e.g., Fractogel materials, have been shown to be very efficient. Similarly, oppositely charged polyarginine and polyglutamate fusion peptides facilitate a specific association of proteins fused with these tags. A thermodynamic quantification of this association reaction was analyzed by fluorescence resonance energy transfer using blue fluorescent protein (BFP) fused to a peptide containing eight glutamic acid residues and enhanced green fluorescent protein (EGFP) with eight C-terminal arginine residues. From a titration of the polyionic EGFP to the oppositely charged polyionic BFP in 25 mM phosphate buffer pH 7.4, a dissociation constant of the polyionic interaction of 250 nM was calculated (Figure 2, Frost and Lilie, unpublished data) translating into an interaction energy of ΔG=-9.4 kcal/mol. If this titration is performed and one of the proteins is not tagged no interaction could be observed, clearly indicating that association is mediated solely via the polyionic fusion peptides.

The affinity of this polyionic interaction strongly depends on the ionic strength of the buffer. At 56 mM NaCl (5 μM BFP and EGFP, respectively, in 25 mM phosphate, pH 7.4) 50% of the BFP-EGFP complex is dissociated; under physiological salt concentrations (150 mM) only 20% of the complex remained (Frost and Lilie, unpublished results). Consequently, polyionic fusion peptides of at

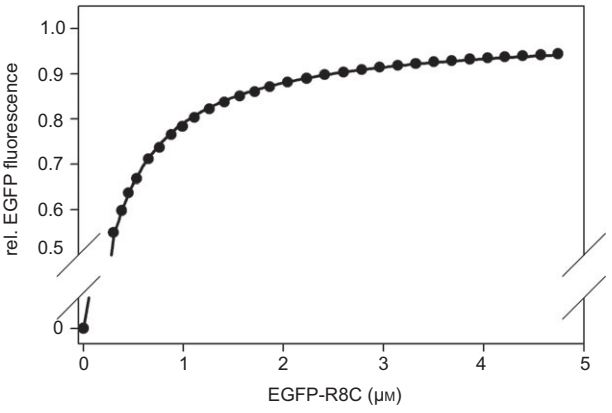

Figure 2 Non-covalent polyionic association of BFP and EGFP. EGFP fused to an octaarginine tag was titrated to 0.3 μM BFP containing a C-terminal octaglutamate tag. The association was monitored in 25 mM phosphate, pH 7.4, 1 mM EDTA, 1 mM DTT by analyzing the EGFP fluorescence at 507 nm after excitation of BFP at 352 nm. The solid line represents a fit to the data according to a dissociation constant of K_D=254 nM.

least eight charged residues facilitate a very specific and structurally highly directed complex formation at low salt conditions *in vitro*. In a biological context such as those *in vivo* this polyionic interaction is not strong enough to ensure stably associated complexes.

Cysteine-containing polyionic fusion peptides induce a covalent association of tagged proteins

The interaction of oppositely charged polyionic fusion peptides is a very specific reaction. The resulting complex, however, is not very stable because of its sensitivity to increasing ionic strength. This sensitivity can be suppressed by introducing cysteines into the polyionic fusion peptides. These cysteines can form covalent disulfide bonds upon association of the polyionic tags, thus rendering the now covalently linked complex insensitive to varying salt concentrations. The advantage of such a system compared to just single cysteines relies on its specificity; whereas single cysteines added at two different proteins will cause covalent association of both homodimers and heterodimers, in the context of polyionic tags covalent heterodimers are formed exclusively because of the specific interaction of the oppositely charged sequences and electrostatic repulsion of identically charged fusion peptides. In a similar approach, Pack and Plückthun (1992) used disulfide-stabilized heterodimerization motifs of four-helix bundles or coiled-coils to design mini-antibodies consisting of covalently linked V_L and V_H domains.

The formation of the disulfide bond between cysteine-containing polyionic fusion peptides is dependent on the redox conditions of the environment. This dependence has been measured by analyzing the amount of covalent heterodimers after incubating synthetic peptides $AlaCysLys_8$ and $AlaCysGlu_8$ in 20 mM borate pH 8.5 in the presence of different ratios of reduced and oxidized glutathione (Figure 3). From this dependency, a redox potential of ΔE^0=-380 mV of the cysteine-containing polyionic peptides could be calculated. Compared to the redox potential of the tripeptide glutathione of ΔE^0=-295 mV under the same conditions (Wunderlich and Glockshuber, 1993), this indicates a very strong interaction of these polyionic peptides, thus favoring the oxidation of the respective disulfide bond compared to that of glutathione (Richter et al., 2001). Similar results were obtained when analyzing the fluorescent proteins BFP-Glu_8Cys and EGFP-Arg_8Cys (Frost and Lilie, unpublished results). Obviously, the polyionic sequences direct covalent heterodimerization only in the presence of low ionic strength; once formed, the stability of the covalent complex is no longer sensitive to high salt concentrations.

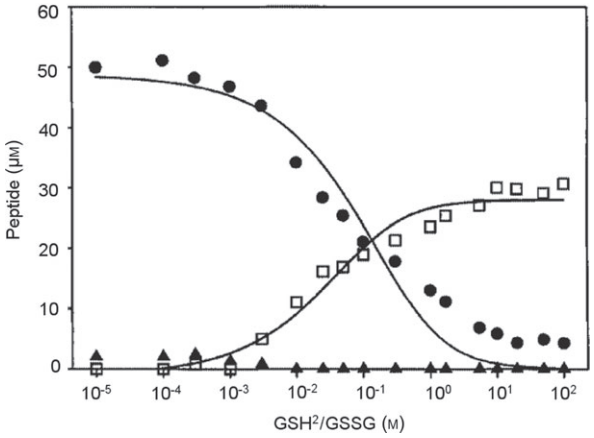

Figure 3 Covalent association of cysteine-containing polyionic peptides.
The covalent heterodimerization of the peptides $AlaCysLys_8$ and $AlaCysGlu_8$ at concentrations of 50 μM was analyzed in dependence of varying concentrations of oxidized and reduced glutathione. Reaction conditions were 20 mM borate, pH 8.5, 20°C. The amounts of covalent heterodimer (closed circle), mixed disulfide $AlaCysLys_8$-glutathione (closed triangle), and free peptide $AlaCysLys_8$ (open squares) were quantified by ion-exchange chromatography. The solid lines represent fits to the law of mass action after reaching equilibrium. From this equilibrium a redox potential of the cysteine-containing polyionic peptides of ΔE=-380 mV could be calculated.

On the experimental level the formation of the disulfide bond leading to a covalent heterodimer is very simple. The oxidation can be carried out just by Cu^{2+}-catalyzed air oxidation of the cysteines, with the specificity of this oxidation being determined by the interaction of the polyionic sequences (Bergelt et al., 2009). For this oxidation, however, it is necessary that the cysteines within the polyionic fusion peptides are in the reduced state prior to the reaction. Alternatively, proteins tagged with cysteine-containing polyionic fusion peptides can be covalently coupled in the presence of a redox system of reduced and oxidized glutathione (Richter et al., 2001; Stubenrauch et al., 2001; Kleinschmidt et al., 2003). The reaction was usually completed within 6–8 h and a product yield of approximately 80% could be obtained (Richter et al., 2001). In one case, however, we observed a drastic limitation in both kinetics and yield of covalent heterodimer formation. The protein granzyme B fused to the peptide $GlyCysAsp_8$ did not covalently couple to an antibody fragment B3-Arg_8CysPro under the conditions mentioned above because the negatively charged peptide $GlyCysAsp_8$ folded back onto a positively charged surface

of granzyme B, thus limiting the accessibility of the tag including its cysteine. Here the covalent coupling of the two proteins could still be obtained by chemically activating the cysteine within the tag of the antibody fragment B3. In order to increase the reactivity of the respective cysteine, it was activated by conversion to a highly reactive mixed disulfide with dithiopyridine (DTP). This activated cysteine then formed a disulfide with the cysteine-containing polyionic fusion peptide of granzyme B; specificity and yield of this reaction were comparable to those mentioned before (Kurschus et al., 2004).

In principle, the cysteine of the polyionic tag can form disulfide bonds with other cysteines of the respective fusion protein as well. This would be especially problematic during oxidative renaturation of proteins linked by a disulfide bond, as the non-native disulfide bond would strongly interfere with formation of the native protein structure upon renaturation. The only example of a disulfide-containing protein tagged with a cysteine-containing octaarginine peptide that has been produced by *in vitro* renaturation is the disulfide-stabilized Fv fragment of the antibody B3 (Stubenrauch et al., 2001). In this case the additional cysteine in the tag did not interfere with the refolding process of the antibody fragment. Presumably, this result cannot be generalized.

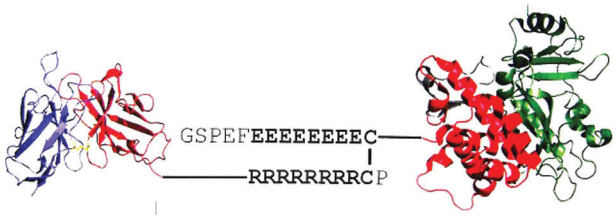

Figure 4 Schematic representation of a polyionic immunotoxin. A disulfide-stabilized antibody Fv fragment (pdb 1DSF) with the V_H domain fused to the polyionic peptide Arg_8CysPro at its C-terminus is covalently coupled by a disulfide bond to the PE38 fragment of *Pseudomonas* exotoxin (pdb 1IKQ) connected to the peptide GlySerProGluPheGlu_8Cys at its N-terminus. The immunotoxin B3-PE38 is listed in Table 1.

Immunotoxins as examples for bifunctional proteins coupled via cysteine-containing polyionic fusion peptides

Cysteine-containing polyionic fusion peptides permit a specific and covalent coupling of two proteins independent of their structure and function. By using common cloning strategies it is thus possible to create a variety of different tailor-made bifunctional proteins. We used this concept for the design of immunotoxins (Figure 4). Immunotoxins are artificially bifunctional proteins consisting of two functional units: (i) a targeting module addressing tumor cells or any other diseased cells; and (ii) a protein exhibiting a toxic activity. The combined functions of these two protein modules result in a cell-type-specific elimination of the targeted cells (Kreitman, 2006). Table 1 lists different immunotoxins designed by the concept of covalent coupling via polyionic fusion

Table 1 Immunotoxins designed by covalent coupling of a targeting module with a protein toxin via polyionic fusion peptides.

Immunotoxin	Targeting module		Toxin		Cytotoxicity [EC_{50} (nM)]	Reference
	Protein	Tag	Protein	Tag		
B3-PE38	Antibody Fv Fragment B3	Arg_8CysPro (C-terminus)	*Pseudomonas* Exotoxin PE38	Glu_8Cys (N-terminus)	0.5	Kleinschmidt et al., 2003
B3-GzmB	Antibody Fv Fragment B3	Arg_8CysPro (C-terminus)	Granzyme B	Asp_8Cys (C-terminus)	35	Kurschus et al., 2004
B3-LLO	Antibody Fv Fragment B3	Arg_8CysPro (C-terminus)	Listeriolysin O	Glu_8Cys (N-terminus)	2.3	Bergelt et al., 2009
CD4-PE38	Domain D1 of Receptor CD4	CysArg_8Pro (C-terminus)	*Pseudomonas* Exotoxin PE38	Glu_8Cys (N-terminus)	10	Jäger and Lilie unpublished
NPY-PE38	NPY peptide	$CysLys_8$ (Coupled to Lys4 side chain of NPY)	*Pseudomonas* Exotoxin PE38	Glu_8Cys (N-terminus)	10	Gehle and Lilie unpublished
B3-VP1	Antibody Fv Fragment B3	Arg_8CysPro (C-terminus)	Polyoma viral Coat VP1	Glu_8Cys (HI-loop)	Gene Delivery	Stubenrauch et al., 2001; May et al., 2002

The cytotoxic activity of immunotoxins depends on the cell lines used in cell culture experiments. The EC_{50} values given for immunotoxins with B3 or NPY as targeting module were measured on MCF7 cells, the CD4-immunotoxin was analyzed on CHO-K1 cells expressing gp120 of HIV.

peptides. In all of these cases, neither the tag nor the coupling affected structure or activity of the single proteins; all polyionic immunotoxins exhibited in cell culture the expected cell-type-specific cytotoxicity with EC_{50} values in the low nanomolar range.

Two other biological activities of polyarginine-tagged proteins have been described. Sequences of seven-to-nine arginine residues mediate a cell-type-unspecific translocation of fused proteins across cellular membranes (Wender et al., 2000; Fuchs and Raines, 2005). Such a cell-penetrating activity of polyarginine has also been used to reprogram human somatic cells using cytokines tagged with eleven arginines (Takahashi et al., 2007). The second known biological activity of polyarginine is the induction of an immune response. Antibodies against a nona-arginine peptide are commercially available, and polylysine with a very similar structure is well known as an efficient adjuvant in immunization protocols (Clementi et al., 1991; Beekman et al., 1999; Wakamoto et al., 2007; Romestand et al., 2010), even though polylysine of a size of c. 120 residues rarely induces an antibody response itself (Romestand et al., 2010). Surprisingly, we observed none of these activities in the case of the polyionically fused immunotoxins or their isolated partner proteins. The cellular interaction of an antibody fragment fused to eight arginines was always dependent on the presence of its cellular antigen, a cell-type-unspecific interaction of the octa-arginine-tag or even a cell-penetrating activity did not occur (May et al., 2002; Kleinschmidt et al., 2003; Bergelt et al., 2009). Presumably, a peptide length of eight arginines is just too short for such an activity. Similarly, a polyionic peptide consisting of eight charged residues might be too small to elicit an antibody response. In classical immunization trials in rabbits using *Pseudomonas* exotoxin, PE38 tagged with an octa-arginine peptide and a disulfide-stabilized B3 Fv fragment fused to an octaglutamate tag, respectively, only antibodies against the respective protein part but not against the polyionic fusion peptide could be raised (Frost and Lilie, unpublished). Thus, for biological applications of artificial bifunctional proteins, such as immunotoxins, polyionic fusion peptides consisting of maximally eight charged residues seem to be well suited as they convey a highly specific coupling reaction of the partner proteins but do not induce any unwanted biological side reactions.

Conclusion

Protein tags are currently very popular in protein chemistry. They are used predominantly in protein purification, however protein tags have several other applications as well. They can help increase expression yields in recombinant protein production, enhance solubility of protein fusions and improve protein stability, especially against proteolytic degradation. Protein tags are indispensible for protein immobilization and can be used in the design of bi- or multifunctional protein constructs because of their ability to induce specific protein-protein interactions. Unfortunately, there is no single tag that is perfectly suited to all of these different applications. As discussed in this review polyionic fusion peptides with or without additional cysteine residues are very versatile in their applications in protein chemistry. Because of the high density of charges in these polyionic sequences, the interaction with oppositely charged polyions either on solid supports or in other proteins is very specific. These interactions can be utilized in protein purification, immobilization, matrix-assisted refolding and the design of artificial bifunctional proteins. Despite the fact that for each single application there are other quite efficient alternatives, such as the His-tag for protein purification or coiled-coil peptides for very specific and extremely stable induced protein association, polyionic fusion peptides are very interesting because of their manifold applications in successive process steps from recombinant protein expression to the design of tailor-made protein complexes.

Acknowledgements: This work was supported by the Deutsche Forschungsgemeinschaft (DFG) as part of the Sonderforschungsbereich SFB 610. We thank Gary Sawers for critically reading the manuscript.

Received January 15, 2013; accepted April 24, 2013; previously published online April 27, 2013

References

Beekman, N.J., Schaaper, W.M., Turkstra, J.A., and Meloen, R.H. (1999). Highly immunogenic and fully synthetic peptide-carrier constructs targetting GnRH. Vaccine *17*, 2043–2050.

Bergelt, S., Frost, S., and Lilie, H. (2009). Lysteriolysin O as a cytotoxic component of an immunotoxin. Prot. Science *18*, 1210–1220.

Bucher, M.H., Evdokimov, A.G., and Waugh, D.S. (2002). Differential effects of short affinity tags on the crystallization of Pyrococcus furiosus maltodextrin-binding protein. Biol. Cryst. *58*, 392–397.

Chao, H., Bautista, D.L., Litowski, J., Irvin, R.T., and Hodges, R.S. (1998). Use of a heterodimeric coiled-coil system for biosensor application and affinity purification. J. Chromatogr. B Biomed. Sci. Appl. *715*, 307–329.

Clementi, M.E., Marini, S., Condò, S.G., and Giardina, B. (1991). Antibodies against small molecules. Ann. Ist. Super. Sanita. *27*, 139–143.

Craig, P.O., Alzogaray, V., and Goldbaum, F.A. (2012). Polymeric Display of Proteins through High Affinity Leucine Zipper Peptide Adaptors. Biomacromolecules *13*, 1112–1121.

Creighton, T.E. (1985). Folding of proteins adsorbed reversibly to ion-exchange resins. In: UCLA Symposia on molecular and cellular biology, Vol. 39, D.L. Oxender, ed. (New York: Academic Press), pp. 249–258.

De Bernardez Clark, E., Schwarz, E., and Rudolph, R. (1999). Inhibition of aggregation side reactions during in vitro protein folding. Methods Enzymol. *309*, 217–236.

Derewenda, Z.S. (2004). The use of recombinant methods and molecular engineering in protein crystallization. Methods *34*, 354–363.

Fuchs, S.M. and Raines, R.T. (2005). Polyarginine as a multifunctional fusion tag. Prot. Science *14*, 1538–1544.

Ford, C.F., Suominen, I., and Glatz, C.E. (1991). Fusion tails for the recovery and purification of recombinant proteins. Protein Expr. Purif. *2*, 95–107.

Hearn, M.T.W. and Acosta, D. (2001). Applications of novel affinity cassette methods: use of peptide fusion handles for the purification of recombinant proteins. J. Mol. Recognit. *14*, 323–369.

Jacob, E. and Unger, R. (2007). A tale of two tails: why are terminal residues of proteins exposed? Bioinformatics *23*, 225–230.

Jana, S. and Deb, J.K. (2005). Strategies for efficient production of heterologous proteins in Escherichia coli. Appl. Microbiol. Biotechnol. *67*, 289–298.

Kiefhaber, T., Rudolph, R., Kohler, H.H., and Buchner, J. (1991). Protein aggregation in vitro and in vivo: a quantitative model of the kinetic competition between folding and aggregation. Biotechnology (N Y) *9*, 825–829.

Kim, S.G., Min, W.K., Rho, Y.T., and Seo, J.H. (2012). Electrostatic interaction-induced inclusion body formation of glucagon-like peptide-1 fused with ubiquitin and cationic tag. Prot. Expr. Purif. *84*, 38–46.

Kleinschmidt, M., Rudolph, R., and Lilie, H. (2003). Design of a modular immunotoxin connected by polyionic adapter peptides. J. Mol. Biol. *327*, 445–452.

Kreitman, R.J. (2006). Immunotoxins for targeted cancer therapy. AAPS J. *8*, 532–351.

Kurschus, F.C., Kleinschmidt, M., Fellows, E., Dornmair, K., Rudolph, R., Lilie, H., and Jenne, D.E. (2004). Killing of target cells by re-directed granzyme B in the absence of perforin. FEBS Lett. *562*, 87–92.

Marston, F.A. (1986). The purification of eukaryotic polypeptides synthesized in Escherichia coli. Biochem J. *240*, 1–12.

May, T., Gleiter, S., Rudolph, R., and Lilie, H. (2002). Assessment of cell type specific gene transfer by polyoma virus like particles presenting a tumor specific antibody Fv fragment. J. Virol. Methods *105*, 147–155.

Nagai, K. and Thogersen, H.C. (1984). Generation of beta-globin by sequence-specific proteolysis of a hybrid protein produced in *Escherichia coli*. Nature *309*, 810–812.

Negro, A., Onisto, M., Grassato, L., Caenazzo, C., and Garbisa, S. (1997). Recombinant human TIMP-3 from *Escherichia coli*: synthesis, refolding, physico-chemical and functional insights. Protein Eng. *10*, 593–599.

Nock, S., Spudich, J.A., and Wagner, P. (1997). Reversible, site-specific immobilization of polyarginine-tagged fusion proteins on mica surfaces. FEBS Lett. *414*, 233–238.

Pack, P. and Plückthun, A. (1992). Miniantibodies: use of amphipathic helices to produce functional, flexibly linked dimeric FV fragments with high avidity in Escherichia coli. Biochemistry *31*, 1579–1584.

Parker, D.E., Glatz, C.E., Ford, C.F., Gendel, S.M., Suominen, I., and Rougvie, M.A. (1990). Recovery of a charged fusion protein from cell extracts by polyelectrolyte precipitation. Biotechnol. Bioeng. *36*, 467–475.

Pastan, I., Hassan, R., FitzGerald, D.J., and Kreitman, R.J. (2007). Immunotoxin treatment of cancer. Annu. Rev. Med. *58*, 221–237.

Richter, S., Stubenrauch, K., Lilie, H., and Rudolph, R. (2001). Polyionic fusion peptides function as specific dimerization motifs. Prot. Engineering *14*, 775–783.

Rogl, H., Kosemund, K., Kuhlbrandt, W., and Collinson, I. (1998). Refolding of *Escherichia coli* produced membrane protein inclusion bodies immobilised by nickel chelating chromaography. FEBS Lett. *432*, 21–26.

Romestand, B., Rolland, J.L., Commeyras, A., Coussot, G., Desvignes, I., Pascal, R., and Vandenabeele-Trambouze, O. (2010). Dendrigraft poly-L-lysine: a non-immunogenic synthetic carrier for antibody production. Biomacromolecules *11*, 1169–1173.

Sassenfeld, H.M. and Brewer, S.J. (1984). A polypeptide fusion designed for the purification of recombinant proteins. Bio/Technology *2*, 76–81.

Schmidt, T.G. and Skerra, A. (2007). The Strep-tag system for one-step purification and high-affinity detection or capturing of proteins. Nat. Protoc. *2*, 1528–1535.

Stempfer, G., Höll-Neugebauer, B., and Rudolph, R. (1996a). Improved refolding of an immobilized fusion protein. Nat. Biotechnol. *14*, 329–334.

Stempfer, G., Höll-Neugebauer, B., Kopetzki, E., and Rudolph, R. (1996b). A fusion protein designed for noncovalent immobilization: stability, enzymatic activity, and use in an enzyme reactor. Nat. Biotechnol. *14*, 481–484.

Stubenrauch, K., Bachmann, A., Rudolph, R., and Lilie, H. (2000). Purification of a viral coat protein by an engineered polyionic sequence. J. Chrom. B *737*, 77–84.

Stubenrauch, K., Gleiter, S., Brinkmann, U., Rudolph, R., and Lilie, H. (2001). Tumor cell specific targeting and gene transfer by recombinant polyoma virus like particle/antibody conjugates. Biochem. J. *356*, 867–873.

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell *131*, 861–872.

Terpe, K. (2003). Overview of tag protein fusion: from molecular and biochemical fundamentals to commercial systems. Appl. Micreobiol. Biotechnol. *60*, 523–533.

Young, C.L., Britton, Z.T., and Robinson, A.S. (2012). Recombinant protein expression and purification: a comprehensive review of affinity tags and microbial applications. Biotechnol. J. *7*, 620–634.

Wakamoto, H., Matsuda, H., Kawamoto, K., and Makino, S. (2007). Epsilon-polylysine microparticle adjuvant drives cytokine production to Th1 profile. J. Vet. Med. Sci. *69*, 717–723.

Weldon, J.E. and Pastan, I. (2011). A guide to taming a toxin–recombinant immunotoxins constructed from Pseudomonas exotoxin A for the treatment of cancer. FEBS J. *278*, 4683–4700.

Wender, P.A., Mitchell, D.J., Pattabiraman, K., Pelkey, E.T., Steinman, L., and Rothbard, J.B. (2000). The design, synthesis, and evaluation of molecules that enable or enhance cellular uptake: Peptoid molecular transporters. Proc. Natl. Acad. Sci. *97*, 13003–13008.

Wunderlich, M. and Glockshuber, R. (1993). Redox properties of protein disulfide isomerase (DsbA) from Escherichia coli. Protein Sci. *2*, 717–26.

Zhao, J.Y., Ford, C.F., Glatz, C.E., Rougvie, M.A., and Gendel, S.M. (1990). Polyelectrolyte precipitation of beta-galactosidase fusions containing poly-aspartic acid tails. J. Biotechnol. *14*, 273–283.

Zouhar, J., Nanak, E., and Brzobohaty, B. (1999). Expression, single-step purification, and matrixassisted refolding of a maize cytokinin glucoside-specific beta-glucosidase. Protein Expr. Purif. *17*, 153–162.