Chemical Science

EDGE ARTICLE

View Article Online
View Journal | View Issue

Cite this: *Chem. Sci.*, 2019, **10**, 427

All publication charges for this article have been paid for by the Royal Society of Chemistry

Received 28th July 2018
Accepted 9th October 2018

DOI: 10.1039/c8sc03355b

rsc.li/chemical-science

Site-selective protein conjugation at histidine†

Karolina Peciak,‡[ab] Emmanuelle Laurine,§[b] Rita Tommasi,¶[b] Ji-won Choi[b] and Steve Brocchini *[a]

Site-selective conjugation generally requires both (i) molecular engineering of the protein of interest to introduce a conjugation site at a defined location and (ii) a site-specific conjugation technology. Three N-terminal interferon α2-a (IFN) variants with truncated histidine tags were prepared and conjugation was examined using a bis-alkylation reagent, $PEG_{(10kDa)}$-mono-sulfone **3**. A histidine tag comprised of two histidines separated by a glycine (His_2-tag) underwent PEGylation. Two more IFN variants were then prepared with the His_2-tag engineered at different locations in IFN. Another IFN variant was prepared with the His-tag introduced in an α-helix, and required three contiguous histidines to ensure that two histidine residues in the correct conformation would be available for conjugation. Since histidine is a natural amino acid, routine methods of site-directed mutagenesis were used to generate the IFN variants from *E. coli* in soluble form at titres comparable to native IFN. PEGylation conversions ranged from 28–39%. A single step purification process gave essentially the pure PEG–IFN variant (>97% by RP-HPLC) in high recovery with isolated yields ranging from 21–33%. The level of retained bioactivity was strongly dependent on the site of PEG conjugation. The highest biological activity of 74% was retained for the PEG_{10}-106(HGHG)-IFN variant which is unprecedented for a PEGylated IFN. The His_2-tag at 106(HGHG)-IFN is engineered at the flexible loop most distant from IFN interaction with its dimeric receptor. The biological activity for the PEG_{10}-5(HGH)-IFN variant was determined to be 17% which is comparable to other PEGylated IFN conjugates achieved at or near the N-terminus that have been previously described. The lowest retained activity (10%) was reported for PEG_{10}-120(HHH)-IFN which was prepared as a negative control targeting a IFN site thought to be involved in receptor binding. The presence of two histidines as a His_2-tag to generate a site-selective target for bis-alkylating PEGylation is a feasible approach for achieving site-selective PEGylation. The use of a His_2-tag to strategically engineer a conjugation site in a protein location can result in maximising the retention of the biological activity following protein modification.

Introduction

The functionality and efficacy of therapeutic proteins can be increased by the covalent conjugation of drugs, probes and polymers (*e.g.* antibody drug conjugates (ADCs) and PEGylation). Most proteins will have regions and sites in their structure where conjugation can result in optimal stability, activity, and pharmacokinetics.[1–3] Computational strategies are being used to identify sites in the protein structure to predict the impact of PEGylation.[1] Site-selective conjugation generally requires both (i) molecular engineering of the protein of interest to introduce a conjugation site at a defined location and (ii) a site-specific conjugation technology.[4]

Many conjugation methods have been described[5,6] and some[7–13] can be specific for an amino acid residue or a glycosyl moiety. Cysteine incorporation into a protein as a site for conjugation has been described in many preclinical studies[14,15] including studies where large numbers of variants have been designed to determine the best site for an unpaired cysteine.[16] Introduction of an additional cysteine can cause scrambling of the native disulfides, and protein dimerisation, oxidation, and aggregation.[17,18] The presence of an accessible cysteine for conjugation can also cause protein aggregation during downstream processing by forming intermolecular disulfides. The unpaired single cysteines engineered into a protein can also be blocked for conjugation by cysteinylation and glutathionylation, which are referred to as 'cysteine capping'.[19] These oxidised forms of the added cysteine must be reduced before conjugation can be conducted, which can often be difficult to accomplish without reducing natural disulfide bonds that may

[a]*UCL School of Pharmacy, University College London, 29-39 Brunswick Square, London, WC1N 1AX, UK. E-mail: steve.brocchini@ucl.ac.uk*

[b]*Abzena, Babraham Research Campus, Babraham, Cambridge CB22 3AT, UK*

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c8sc03355b

‡ Current address: Ipsen Biopharm Limited, Ash Road, Wrexham Industrial Estate, Wrexham, LL13 9UF, UK.

§ Current address: Bio-LogicIP, 1 Westbourne Road, N7 8AR London, UK.

¶ Current address: Lonza Biologics, 228 Bath Road, SL1 4DX Slough, UK.

be present in the protein. Incorporation of an unpaired cysteine for conjugation is best utilised in proteins that do not have a native disulfide.

Site-selective conjugation can be achieved by the incorporation of noncanonical amino acids (ncAA), which are sometimes referred to as non-natural amino acids (NNAAs).[20–25] As with cysteine incorporation, ncAA incorporation can be achieved by addition or substitution of a single amino acid into the protein of interest. ncAA incorporation is typically achieved by genetically reassigning a stop codon as a recognition site for an engineered pair of tRNA and aminoacyl-tRNA synthetase that incorporates the ncAA at the stop codon.[26] However, protein expression using stop codon reassignment can result in "leak expression"[27] and low yields of protein expression[28] due to low overall suppression efficiency defined as the percentage of stop codon read-through, and due to mechanisms that degrade mRNA[29] containing premature stop codons, called nonsense mediated mRNA decay. Strategies have been described[1,26,30] such as using cell-free protein synthesis to address limitations[31] for the specific incorporation of ncAAs.

ncAAs have side chains (*e.g.* keto, azide or alkyne) with orthogonal reactivity to native amino acids.[20–25] Conjugation conversions and yields with ncAA modified proteins can be high[32] depending on many factors including reagent stoichiometry and the presence of catalysts (*e.g.* copper). Orthogonal reaction conditions must be suitable for protein modification. For example, the use of copper in click approaches can be challenging due to toxicity[33] and can cause protein misfolding, aggregation and degradation.[34] Copper readily complexes to proteins and is the basis for quantification assays, so it is not clear why there has been so much effort to use copper in conjugation reactions with proteins.

Strain-promoted azide alkyne cycloadditions (SPACC) is one alternative to avoiding copper. However, hydrophobicity caused by conjugation linker moieties[35] or due to surface accessible ncAAs can cause non-specific protein binding.[36] Although the impact of hydrophobicity will be less at the site of conjugation for PEGylation applications, ADC hydrophobicity has also long been recognised to cause accelerated clearance and reduced efficacy.[37,38] ADC hydrophobicity is also caused by linker–drug structure and the drug–antibody ratio (*i.e.* drug loading). Strategies have been described to solve the challenges caused by hydrophobicity to (i) reduce non-specific binding in azide alkyne cycloadditions[39,40] and (ii) optimise ADC development.[41,42]

Increased conjugation selectivity can also be accomplished by removal of amino acid residues that undergo competitive conjugation. This strategy was followed for the development of PEGvisomant,[43] which is a growth hormone antagonist. Lysine residues were substituted to arginine to remove PEGylation sites that caused unacceptable reduction in biological activity. Lysine depletion approaches have been used to promote N-terminal[44] or to have a single lysine residue remaining to promote site-selected conjugation.[44–46] Use of lysine depletion methods is challenging because lysines are usually abundant[47] and are necessary for protein activity.[45,48–50] Strategies continue to be developed for lysine specific conjugation.[51]

We have found that the PEG-bis-sulfones **1** can undergo site-specific bis-alkylation with the two thiols from a native disulfide bond in a protein (Fig. 1A and B).[52,53] The bis-sulfone **1** undergoes elimination of a toluene sulfinic acid (*e.g.* **2**) to the mono-sulfone **3** which is then capable of undergoing a sequence of addition–elimination reactions to site-specifically modify the protein by re-bridging a reduced disulfide with a three carbon bridge attached to PEG. Other bridging reagents have been described,[54–57] and we have continued to examine reagents based on the PEG-bis-sulfone **1** due to the simplicity of their preparation and their practical application to prepare ADCs[58–60] and antibody mimetics.[61–63]

We have also found that the PEG bis- and mono-sulfones (**1** and **3**) can site-specifically undergo conjugation to the histidines in a C- or N-terminal histidine tag (His-tag) (Fig. 1C)[64] without need of metal chelation. His-tags are a contiguous sequence of histidines (*e.g.* six histidines, His_6-tag), which are widely used in a broad range of applications in protein science and enable effective purification and can also increase protein expression yields.[65,66] His-tags are usually placed at a C- or N-terminus of a protein where often one or both of these regions of a protein do not adversely compete with protein function, *e.g.* binding or activity. Our reasoning was that PEGylation at a terminal His-tag would display higher activity, which indeed is what was observed with interferon α2-a and a domain antibody.[64] Another way to target His-tags for conjugation is to use reagents chelated to nickel or zinc which can then complex to the imidazoles in a His-tag on a protein. Applications tend to be for protein labeling and immobilisation.[67–69] Covalent bonds can be formed after complexation by subsequent reaction by other functionality on the reagent.[70,71]

During our studies of His-tag PEGylation,[64] preliminary experiments were also conducted with a small peptide called chemerin-9 peptide (149-YFPGQFAFS-157). Varying numbers of histidines (*e.g.* two, four or six) and a histidine–glycine tag (**HGHGHG**) were appended to the C-terminus of the peptide. Chemerin-9 peptide does not contain any lysines and PEGylation was only observed when histidines were added to the peptide. We now examine the site-selective conjugation of PEG-mono-sulfone **3** to a truncated histidine tag comprised of 2 histidines (His_2-tag) engineered into a protein that also has several lysine residues. Compared to a 6 histidine tag (His_6-tag), a His_2-tag could be positioned at more locations within the protein for site-selective conjugation while ensuring protein function can be maintained.

Most therapeutic proteins do not have contiguous histidines in their sequence, although some proteins are known to have clusters of histidines at their surface.[72] While histidine is a relatively rare amino acid in proteins[73] using bis-alkylation to target two histidines that are close together rather than a strategy to target one histidine will avoid the problem of competitive conjugation to other histidines in the protein of interest. There are three natural histidines and 11 lysine residues in interferon α2, none of which were conjugated when there was a N-terminal His-tag on interferon α2.[64]

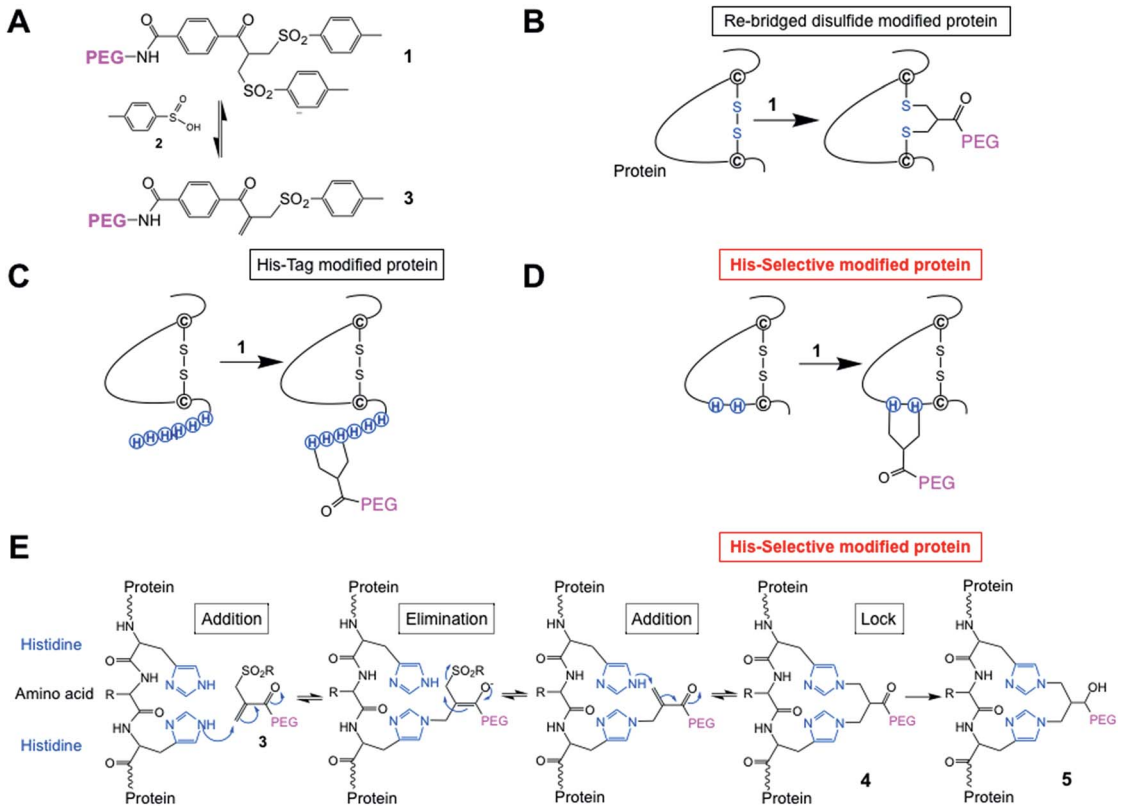


Fig. 1 (A) Chemical structure of the PEG-bis-sulfone **1** and formation of the PEG-mono-sulfone reagent **3** generated after elimination of toluene sulfinic acid **2**; (B) both PEG-bis- and mono-sulfone reagents (**1** and **3**) can undergo site-specific bis-alkylation with the two cysteine thiols from a native disulfide in a protein achieving disulfide-bridging PEGylation;[52] (C) both reagents (**1** and **3**) also undergo conjugation to the histidines in a C- or N-terminal histidine tag (His-tag);[64] (D) herein we investigated site-selective bis-alkylation conjugation at a His_2-tag engineered at a selected site along the protein mainchain; (E) mechanism for conjugation by bis-alkylation at a His–Gly–His tag by a sequence of addition–elimination reactions to the initial conjugate **4**. Carbonyl reduction to give **5** stops de-PEGylation by preventing retro-Michael reactions.

Histidines have a lower pK_a than other nucleophilic residues in a protein, *i.e.*, lysine and arginine, so at mildly acidic pH (5–6) they may not be protonated, so can be reactive. Conjugation by bis-alkylation using reagents such as **1** and **3** can be equilibrium controlled *via* an addition–elimination reaction that is reliant on the Michael reaction. Covalent conjugation occurs with 2 amino acids close in space (*e.g.* the two cysteine thiols from a disulfide[53] or two histidines in a C- or N-terminal his-tag[64]). Conjugation to two lysine amino nucleophiles in slightly acidic conditions is not favoured for reagents **1** and **3** when there are cysteine thiols or histidine imidazoles present. Therefore, using 2 closely placed histidines in interferon α2 (Fig. 1D and E) may provide an alternative method for achieving site-selective conjugation compared to, for example, (i) adding an unpaired cysteine to a protein with existing disulfides and (ii) incorporating ncAAs, which have been described for interferon α2.[15,74]

The aim of the study described herein was to determine if site-selected bis-alkylation conjugation could be accomplished with a two histidine tag (His_2-tag) placed within the primary sequence of interferon α2-a to give a PEGylated interferon with a higher biological activity than has been observed previously for PEGylated interferon α2.

Results and discussion

Preparation of truncated N-terminal His-tag IFN analogues

Interferons are a group of naturally occurring cytokines produced in vertebrates in response to a viral infection.[75] Cytokines are a broad group of proteins that have clinical relevance,[76,77] but quickly clear from circulation, so several cytokines that have been approved for clinical use have also been PEGylated.

Interferons are a key component of the innate immune response and are used clinically to treat a wide range of conditions including viral infections, malignancies and multiple sclerosis. There are three types of interferon (I, II and III), and type-I interferons that are used clinically include interferon α and β. There are at least 13 different human interferon α proteins, all which have 166 amino acids except for interferon α2 which has 165 amino acids due to a deletion at position 44. Like many cytokines, interferon is a helical bundle protein that has a cluster of five α-helices, four of which are arranged to form a left handed helix bundle motif (helices A, B, C and E).[78–80] Between helix A and B is a loop of 30 residues.

Initial experiments were conducted to determine if less than eight histidine residues at the N-terminus of interferon

α2-a could undergo conjugation with PEG_{10}-mono-sulfone **3** in conditions where native interferon α2-a was not PEGylated. The reagent **3** was derived from a PEG with a molecular weight of 10 kDa, so is abbreviated PEG_{10}-mono-sulfone **3**. Interferon α2-a will herein be abbreviated to 'IFN'.

It was first necessary to determine if a truncated N-terminal His_2-tag IFN could undergo conjugation in conditions where native IFN did not undergo conjugation. Selective conjugation was thought more probable to occur in slightly acidic conditions to avoid lysine conjugation, so reactions were conducted with the PEG_{10}-mono-sulfone **3** rather than PEG_{10}-bis-sulfone **1**. The use of the PEG mono-sulfone **3** undergoes conjugation directly without the need for elimination of the first molecule of toluene sulfinic acid **2** (Fig. 1A). The elimination of toluene sulfinic acid **2** is much slower at acidic pH values than at neutral and basic pH values.[53]

Although it was not clear which IFN variants would be expressed, we sought to prepare four different histidine-truncated tags (ESI†). The first two IFN variants were fused with two or four contiguous histidines followed by glycines. Tag 1 (**HH**GGGG-IFN) and Tag 2 (**HHHH**GG-IFN) (Fig. 2A) were followed by glycines to make up a 6-amino acid tag analogous to a typical His_6-tag commonly used in protein expression studies. The remaining two N-terminal IFN variants were prepared with a glycine separating the histidine residues: Tag 3 (**HGH**GGGG-IFN) and Tag 4 (**HGHGH**G-IFN) (Fig. 2A). Flexible protein inter-domains are rich in glycine residues.[81] Our premise was that a His_2-tag separated by a glycine would result in the flexibility needed to ensure the histidines would be positioned to achieve selective and efficient conjugation.

Initial expression experiments were not successful using the pET1010/D-TOPO® vector which had been used to prepare N-terminal 8-histidine tagged IFN ($\mathbf{H_8}$-IFN).[64] We then examined IFN variants that were expressed as a fusion protein with the SUMO-tag using the Champion™ pET SUMO Expression system previously used to produce consensus interferon.[82] The SUMO fusion of the four N-terminally tagged IFN variants were expressed on a 0.5 L scale using the SHuffle® *E. coli* expression strain.

The SUMO fusions of the N-terminally tagged IFN variants were expressed successfully in soluble form and were present in the supernatant fraction (ESI Fig. S1†). The SUMO fusion partner contains a His_6-tag that allows affinity purification of the fusion protein prior to digestion needed to release the desired his-tagged IFN variant. Each SUMO–IFN variant was thus purified from the soluble fraction by immobilised metal affinity chromatography (IMAC) followed by digestion to remove the SUMO fusion partner (Fig. 2B). A second IMAC purification step was used as a negative purification step to separate the cleaved SUMO partner and yield the purified IFN variants (Fig. 2C). Purity of isolated IFN was confirmed by RP-HPLC (Fig. 2D, RP-HPLC are shown in the ESI†).

The Tag 2 variant (**HHHH**GG-IFN) with four contiguous histidine residues was biologically active (Table 1) but was difficult to purify at high enough scale from the SUMO protein, which contains six contiguous histidine residues. Since the aim of these experiments was to select a His_2-tag if possible, it was decided that initial conjugation studies would be conducted using the other three His-tagged IFN variants: (i) Tag 1 (**HH**GGGG-IFN), (ii) Tag 3 (**HGH**GGGG-IFN) and (iii) Tag 4 (**HGHGH**G-IFN). These three IFN variants all displayed comparable activity to the interferon α-2a NIBSC standard in the A549/EMCV antiviral *in vitro* biological assay (Fig. 2E, Table 1).

Conjugation to the truncated N-terminal His-tag IFN variants

PEG conjugation studies with the three His-tagged IFN variants (Tags 1, 3 and 4) were conducted using PEG_{10}-mono-sulfone **3**.

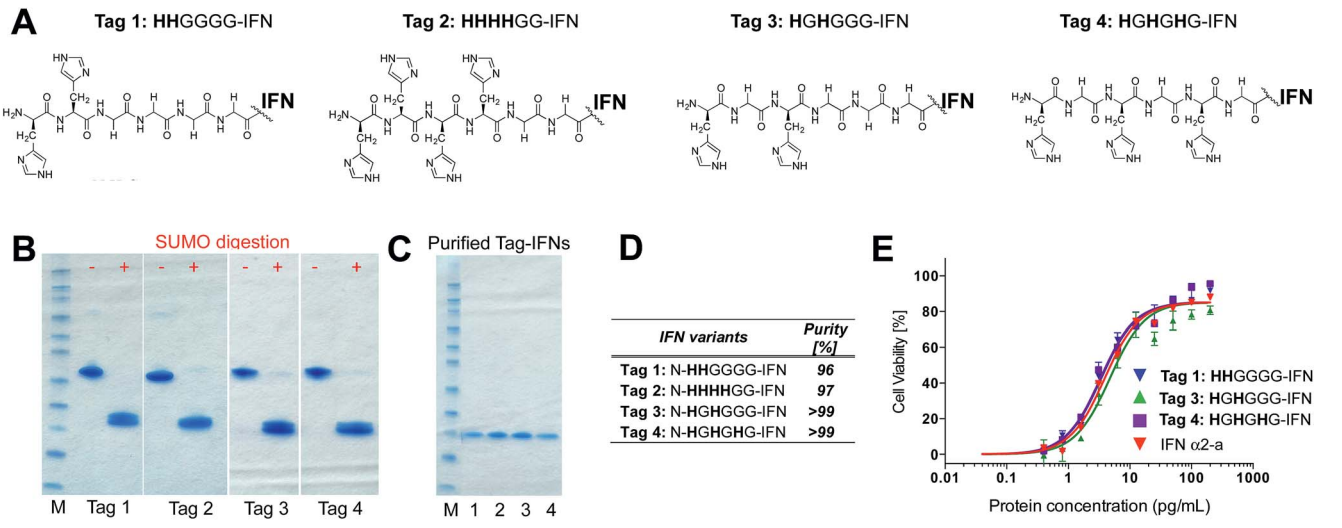


Fig. 2 (A) Four N-terminal histidine tags were designed and fused to IFN to examine if different configurations of histidine residues would undergo selective conjugation with PEG_{10}-mono-sulfone **3**; (B) SDS–PAGE gels (colloidal blue) of the N-terminally histidine tagged IFN–SUMO fusion proteins that were isolated and then digested to separate the SUMO protein from the IFN variant; (C) SDS–PAGE gel (colloidal blue) of the purified N-terminal IFN variants; (D) purities of the N-terminal IFN variants as determined by RP-HPLC. Only small amounts of the Tag 2 (**HHHH**GG-IFN) variant could be isolated, so was not used in the conjugation studies. (E) The three N-terminally tagged IFN variants (Tags 1, 3 and 4) used for conjugation studies displayed equivalent activity compared to native IFN.

Table 1 Antiviral activity of N-terminally tagged IFN variants confirmed that fusion of a histidine–glycine tag to the N-terminus does not impact on IFN activity

	Specific activity calculated to NIBSC IFNα-2a [MIU per mg]				
		Tag 1	Tag 2	Tag 3	Tag 4
Sample	IFN	**HH**GGGG-IFN	**HHHH**GG-IFN	**HGH**GGG-IFN	**HGHGH**G-IFN
n	$n = 3$	$n = 5$	$n = 5$	$n = 5$	$n = 5$
MEAN	254.2	228.5	216.7	242.4	218.4
SEM	34.5	23.2	18.5	31.4	51.4

The 10 kDa molecular weight PEG is comparable to that used in the clinic for PEG-Intron® (12.5 kDa PEG) and many prior studies indicate that a 10 kDa PEG was sufficiently large enough[83] to surpass the molecular weight threshold needed to sterically shield the IFN to reduce its biological activity.

Native IFN[64] was used as a control for these conjugation studies. When the conjugation reaction was conducted at pH 5.0 at a protein concentration of 1 mg mL^{-1} and 5 eq. of PEG$_{10}$-mono-sulfone **3** for an incubation period of 16 h at 20 °C, the conversion to the mono-PEGylated species ranged from 29% to 35% for the three IFN variants (Tags 1, 3 and 4) whilst only 7% PEGylation was observed for the non-His-tagged IFN (Fig. 3A and B).

The **HGHGH**G-IFN variant (Tag 4) with three histidine residues did show an increased tendency to form di- and tri-PEGylated products when 5 equivalents of PEG$_{10}$-mono-sulfone **3** was used in these conditions (*i.e.* incubated for 16 h at 20 °C). Multiple PEGylation is possible for the **HGHGH**G-IFN variant between the three histidines and the terminal amine. The **HH**GGGG and the **HGH**GGG-IFN variants with two histidine residues generated little or no di-PEGylated product. Increased protein concentration and PEG stoichiometry yielded unacceptable amounts of multi-PEG conjugates (Fig. 3A).

Previously with N-terminal octahistidine tagged IFN (**H$_8$**-IFN)[64] we found no evidence of PEG conjugation except on the His-tag suggesting histidine conjugation was faster at the optimal conditions than at other sites on the protein, and that once PEG was conjugated to the His-tag then it is possible that steric effects slowed the rate of non-specific conjugation.

Most optimised conjugation reactions require well defined and mild reaction conditions that are conducive to maintaining protein structure while achieving efficient and selective conjugation. At commercial scale, the cost of the GMP conjugation reagent is comparable or more expensive than the cost of the protein. Scalability considerations require that reagent stoichiometry, the amount of unmodified protein remaining after conjugation and the purification process used to isolate the desired conjugate should all be taken into account. Isolation of the desired conjugate from excess unmodified protein and conjugation reagent can generally be conducted effectively at scale.

Reductive amination strategies which require imine formation followed by mild reduction aim to exploit a small pK_a difference of the N-terminal amine for conjugation compared to the other amines in the protein. For PEG to be predominantly conjugated at the N-terminal amine,[84] a narrow set of

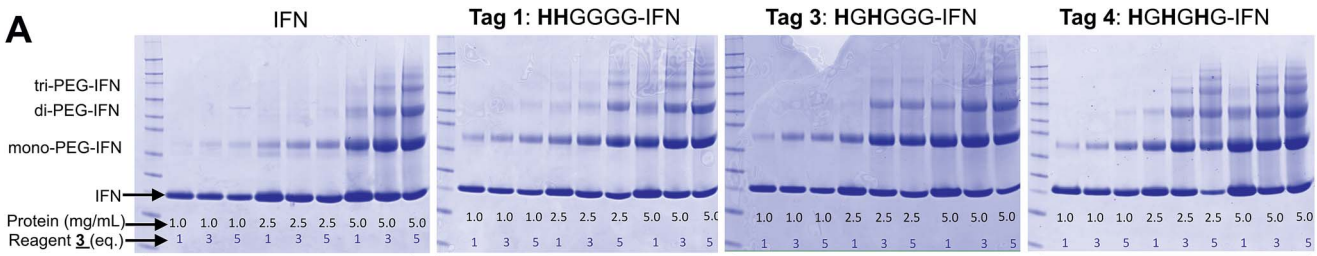


Fig. 3 PEG conjugation studies of truncated N-terminal His-tagged IFN variants (Tags 1, 3 and 4) and native IFN used as control. Protein to PEG$_{10}$-mono-sulfone **3** ratios were screened and the conjugation reaction was conducted for 16 h at 20 °C. It was observed that the most effective conjugation at the histidine-tag was achieved when reaction was conducted at pH 5.0 at a protein concentration of 1 mg mL^{-1} and 5 eq. of PEG$_{10}$-mono-sulfone **3**. The conversion to the mono-PEGylated species ranged from 29–35% for the three IFN variants, while only 7% PEGylation on native IFN. (A) SDS–PAGE gels (colloidal blue) showing the PEG-conjugation reaction mixtures screened and (B) estimates of PEGylation conversion rates using densitometry analysis for the conjugation conducted with protein concentrations of 1.0 and 2.5 mg mL^{-1}.

conditions that are protein dependent is required. Exploiting a slight difference in the pK_a of a single N-terminal amino acid to achieve high yielding and homogeneous imine formation in an aqueous medium is not possible with low reagent stoichiometry or complete site specificity.[85,86] Several strategies have been described to direct conjugation to the N-terminus of a protein to address the inherent limitations of reductive amination using PEG aldehyde reagents.[13,64,87]

The data with the truncated N-terminal histidine tagged IFN variants (Fig. 3A and B) indicate that the best mono-PEGylation conditions were to incubate PEG_{10}-mono-sulfone **3** (5 eq.) with the IFN variant at 1 mg mL^{-1} for 16 h at 20 °C (pH 5.0). Using these conditions, all three of the IFN histidine variants resulted in conjugation broadly comparable to that previously observed with an N-terminal $\mathbf{H_8}$-IFN.[64] Purification of the N-terminal PEG–IFN conjugate was also easy to conduct. A single pair of histidines separated by a glycine, *e.g.* **HGH**, was then incorporated within the IFN mainchain as a site-selective conjugation location.

Site selection on IFN for histidine conjugation

Sites to incorporate an internal His-tag for conjugation were considered based on the published findings of IFN binding interactions with its dimeric receptor and the *in vitro* biological activities of isolated PEG-positional isomers.[88,89] IFN exerts its biological activity through binding to the specific cell surface receptors.[90] Features of the IFN structure and its interaction with cell membrane receptors have been described and key residues mediating these interactions have been determined.[91] Type 1 IFNs bind to a heterodimeric cell surface receptor (IFNAR1 and IFNAR2).[91–95] The amino acid sequences 29–35, 78–95 and 123–140 are thought to bind to IFNAR1 and IFNAR2.[91,94].

The two main clinically used PEGylated interferon products (PEG-Intron® and Pegasys®) are produced with PEGylation reagents that undergo non-specific conjugation on interferon α2 proteins that differ by one amino acid. Pegasys® is produced from interferon α2-a (lysine at position 23) and PEG-Intron® is produced from interferon α2-b (arginine at position 23).

PEG-Intron® is a 12 kDa PEG conjugate of interferon α2-b derived from a linear PEGylation reagent functionalised with an *N*-hydroxy-succinimide (NHS) active ester that undergoes non-selective acylation with protein nucleophiles (*e.g.* terminal amine, lysine amines). There are at least 13 mono-PEGylated interferon α2-b positional isomers in PEG-Intron® with PEG conjugated at lysine, serine and histidine residues, and at the N terminus (Cys1).[88,96] Most of the observed activity is from the positional isomer where PEG is conjugated to His34 (or H34).[97] Reaction of an NHS ester with the histidine imidazole side chain results in the formation of a hydrolytically labile carbonyl imidazolide. The H34 positional isomer is the most hydrolytically labile PEG positional isomer,[97] so the observed activity of 37% for H34 positional isomer is likely due to the presence of de-PEGylated, unmodified interferon α2-b. The *in vitro* antiviral activity of the positional isomer at the amine terminus is ~13%, however the activities of the other positional isomers that have been described are much lower (<10%).

Pegasys® is made using a lysine-derived branched PEG reagent with two 20 kDa PEG molecules linked *via* urethanes to the two lysine amines. An NHS active ester at the lysine carboxylate functions as the protein conjugating moiety which undergoes non-selective acylation with protein nucleophiles. Hydrolysis of NHS functionalised PEG reagents is reported to be competitive at pH values where amine conjugation can occur (*e.g.* pH 7.5–8.0).[98] While an excess of 5 equivalents has been claimed to give a 45% conversion to mono-PEGylated IFN α-2a,[99] other studies indicate the need to use a greater excess of NHS PEGylation reagents (10 eq.).[100] Pegasys® comprises at least nine bioactive positional isomers, with four major isomers PEGylated at the lysine residues at positions 31, 21, 131 and 134.[89,101,102] The *in vitro* activities of the positional isomers range from ~2% to 12%.

Five different IFN variants with internal His-tags (or PEG-tags) were thus prepared (Fig. 4, Table 2, ESI†). The 5(**HGH**)-IFN variant was designed by engineering Q5H and T6G substitutions. One of the three natural histidine residues present within IFN is located at position 7, therefore this amino acid was not substituted. The resulting 5(**HGH**) conjugation tag

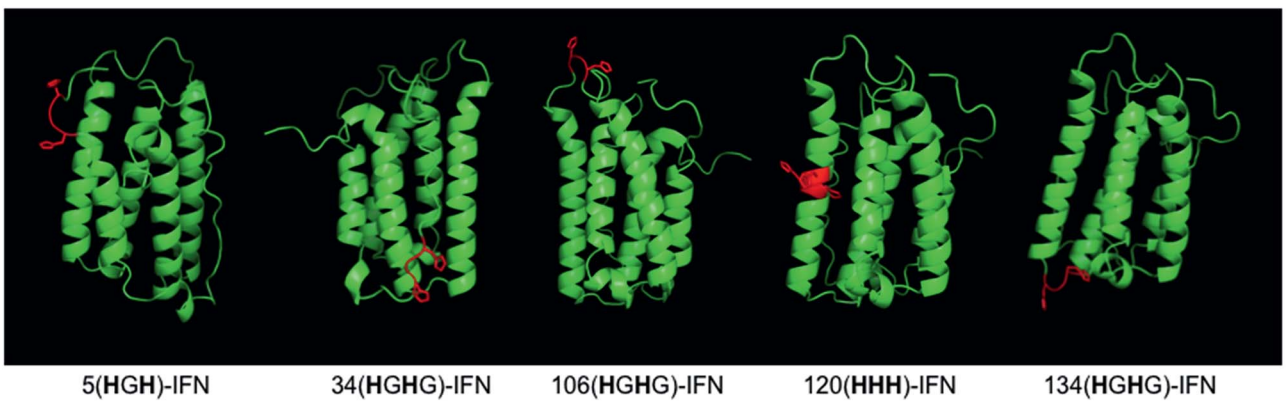


Fig. 4 The IFN variants with an incorporated histidine conjugation tag (shown in red) were rationally designed based on the published findings on IFN binding interactions with its receptor and *in vivo* biological activities of the isolated PEG-positional isomers from marketed products.

Table 2 Detailed description of the internal His-tagged IFN variants showing the localisation of His-tag insertion, sequence of the His-tag and native amino acids that were replaced

IFN	PEG-tag location	His-tag sequence	Original IFN amino acids
5(**HGH**)-IFN	5	**HGH**	QT**H**
34(**HGHG**)-IFN	34	**HGHG**	**H**DFG
106(**HGHG**)-IFN	106	**HGHG**	T-ET
120(**HHH**)-IFN	120	**HHH**	RKY
134(**HGHG**)-IFN	134	**HGHG**	K-YS

is located within a short, flexible and solvent accessible loop in close proximity to the N-terminus of the protein. PEGylation of 5(**HGH**)-IFN allows comparison with the N-terminal **H**$_8$-IFN conjugates previously described[64] and the N-terminal PEG-Intron® positional isomer.

The 34(**HGH**)-IFN variant was prepared to evaluate if we could conjugate a PEG at the His34 (H34) site to compare with PEG-Intron®. The tag was incorporated through D35G and F36H substitutions. H34 is located on the AB loop and is thought to be involved in binding to the IFN receptor, thus it was anticipated that significant reduction of biological activity of native IFN would be observed.

The 106(**HGHG**)-IFN variant was prepared as this position is located within a flexible loop that is thought not to be involved with receptor binding. While glycosylation is rare among native human interferons, it is known that T106 can be O-glycosylated in interferon α2-b.[103,104]

Although IFN has two disulfides (C1–C98 and C29–C138), a study was conducted to insert a free cysteine at position 111 which is at the beginning of the D α-helix after the proline residue in this region.[15] This cysteine IFN variant was PEGylated with maleimide PEG reagents and displayed comparable *in vitro* activity to PEG-Intron® in the Daudi cell growth inhibition assay. There are no reported PEG positional isomers at positions 99–109 in PEG-Intron® or PEGasys® due to the absence of nucleophilic amino acids in this region of the protein.[88,89]

The 106(**HGHG**)-IFN variant was created by making T106H, E107G and T108H substitutions. An additional G108 was inserted to allow more flexibility within the conjugation tag because there is a following proline in the IFN sequence. Proline restricts free rotation and typically appears on the surface of proteins while introducing a β-turn in the amino acid sequence.

Another IFN variant that was prepared is 120(**HHH**)-IFN, which was engineered with R120H, K121H and Y122H substitutions in the D α-helix. This tag is present in the D-helix which is solvent-exposed, but is less flexible compared to the loop regions where other IFN tags are located. The PEG positional isomer at lysine 121 (K121) in PEG-Intron® and PEGasys® displays <10% activity,[88,89] so the 120(**HHH**)-IFN variant was prepared as a negative control. Three contiguous histidine residues (**HHH**) were used in this tag rather than a **HGH** tag to account for the lack of flexibility in the D-helix and to allow the imidazole rings to adopt a conformation for bis-alkylation conjugation.

The final tagged IFN variant, 134(**HGHG**)-IFN, was created using K134H, Y135G and S136H substitutions and insertion of additional glycine. The additional glycine was inserted because of the following proline residue in the sequence. The 134(**HGHG**) tag was created as another negative control to compare against the K134 positional isomer described during the development of Pegasys®,[89] which had displayed low activity (<10%; 1358 ± 46 U μg^{-1}).

Expression of the five internal His-tagged IFN variants was conducted using the SUMO system using SHuffle® *E. coli* strain (ESI†). No soluble expression was observed for 134(**HGH**)-IFN and this variant appeared to be present as insoluble protein aggregates at the top of SDS–PAGE well (Fig. S2†). This variant contains a tag in the binding region of IFN and was not expected to display increased *in vitro* antiviral activity once PEGylated over the other PEGylated IFN positional isomers described herein. No further attempts were made to express this protein.

Fermentations (1 L) were conducted with the four remaining variants to give 90–110 mg of the SUMO–IFN protein per variant (ESI†). Each IFN variant was obtained after SUMO digestion and isolated by a sequential IMAC–anion exchange purification process. IMAC retained the SUMO fusion partner and the SUMO protease while allowing the elution of the IFN variants onto an anionic ion exchange chromatography (IEC) column. Each bound IFN variant was eluted at a concentration of ~1 mg mL^{-1}

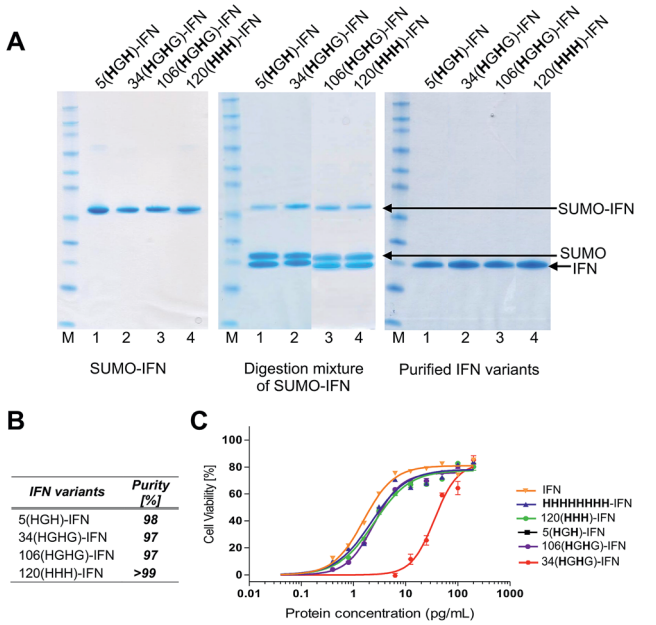


Fig. 5 (A) SDS–PAGE gels (colloidal blue) of the five internal His-tagged IFN variants which had been expressed as IFN–SUMO-fusion proteins (colloidal blue detection). Proteolytic digestion with SUMO protease followed by IMAC purification yielded the designed IFN variants; (B) all of the histidine IFN variants were purified to high purity as determined by RP-HPLC; (C) antiviral activity showed that three histidine IFN variants had similar activity to native IFN, whereas the 34(HGH)-IFN variant was low (~5%) compared to native IFN. The reduction in antiviral activity was not unexpected as the natural histidine located in position 34 in IFN is known to be important for IFN biological activity.

to be used for conjugation studies (Fig. 5A) (ESI†). The four IFN variants were isolated in good purity (Fig. 5B, RP-HPLC are shown in the ESI†).

The process of protein engineering and verification of protein expression was shown to be straightforward and took less than two weeks. All of the expressed IFN variants remained soluble and stable following removal of the SUMO fusion partner. The expression method that was used to obtain the IFN variants did not require further optimisation from the method established for expression of native IFN.[82] The *in vitro* antiviral activity of the four internal His-tagged IFN variants was determined using the A549/EMCV antiviral assay (Fig. 5C) using the NIBSC standard for IFN as control. The specific activities of the 5(**HGH**)-IFN, 106(**HGHG**)-IFN and 120(**HHH**)-IFN variants were similar to the NIBSC standard (Fig. 5C). The activity of the 34(**HGH**)-IFN variant was considerably lower (~5% compared to native IFN) so this variant was not used in subsequent conjugation studies. The 34(**HGH**)-IFN variant was produced to compare with the H34 PEG positional isomer in PEG-Intron®.[97] The reduced activity of 34(**HGH**)-IFN was not unexpected as the natural histidine located in position 34 in IFN is known to be important for biological activity.[96,105–107]

Site-selective conjugation studies

Conjugation with PEG_{10}-mono-sulfone **3** was examined with the three remaining IFN variants (Fig. 6A) using the conditions determined previously for the N-terminal truncated His-tagged IFNs (Fig. 3). The conjugation reactions were conducted with 5 molar equivalents of PEG_{10}-mono-sulfone **3** and allowed to incubate overnight at 20 °C followed by analysis using SDS–PAGE (Fig. 6B). Conversions to the mono-PEGylated conjugates (28–39%) were similar to conversions observed during conjugation of the N-terminally His-tagged IFN variants (Fig. 3B).[64] The stoichiometry of the PEG_{10}-mono-sulfone **3** and the conversions to the PEG–IFN conjugate also compare favourably to other amine conjugation methods.[99]

Trace di-PEGylated product was sometimes observed by SDS–PAGE and was readily removed during purification. Non-specific PEGylation may have occurred in addition to PEGylation on the

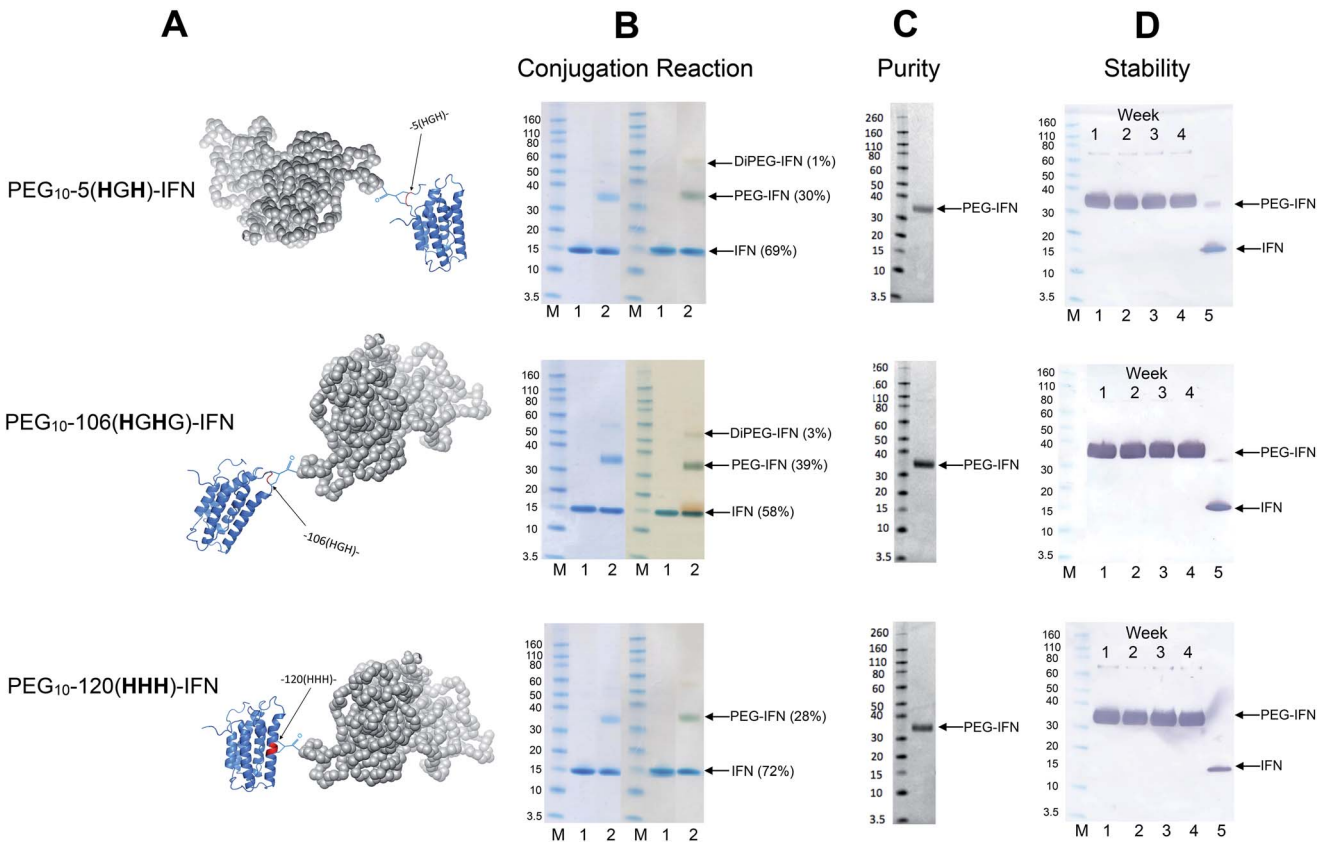


Fig. 6 PEG conjugation studies were conducted with three of the internal His-tagged IFN variants (shown in A). PEG_{10}-mono-sulfone **3** (5 eq.) was used for conjugation with each IFN variant at 1 mg mL^{-1}. The conjugation reaction was conducted for 16 h at 20 °C and the observed conjugation conversions ranged between 28–39% and were similar to the N-terminally tagged IFNs. (B) SDS–PAGE gels for each IFN variant conjugation: lane M protein markers, lane 1, the IFN variant used for PEGylation and lane 2, the PEGylation reaction mixture. For each IFN variant, the left gel is stained with colloidal blue and the right gel is further stained with barium iodide. (C) Purified mono-PEGylated IFN variants were isolated from the crude conjugation reaction mixture using a single ion exchange purification step to deliver a final product in high purity as determined by SDS–PAGE (silver stain) and by RP-HPLC (Fig. S3†); (D) stability studies were conducted for 4 weeks at 4 °C in 50 mM sodium phosphate, 150 mM NaCl, pH 7.4 buffer with each PEG–IFN conjugate at 0.2 mg mL^{-1}. All of the three site-selectively PEGylated IFN variants were stable to de-PEGylation with no evidence of free IFN being detected with anti-IFN western blot analysis, lane M protein markers, lane 1, week 1, lane 2, week 2, lane 3, week 3, lane 4, week 4 and lane 5, native IFN used as control.

His-tag. Prior to purification, the PEGylated IFN variants were treated with sodium triacetoxyborohydride to reduce the aryl ketone to prevent retro-Michael induced PEG deconjugation.[64,108] Carbonyl reduction using a mild hydride reagent acts as a stabilisation step and is comparable to the analogous process conducted during reductive amination[85,86,109–111] in the manufacture of pegfilgrastim (Neulasta®).[84] It has become evident even with maleimides that it is necessary to conduct a post-conjugation hydrolysis stabilisation step to avoid deconjugation.[112–114]

The mono-PEG conjugates were isolated from the crude reaction mixtures by a single IEC step. Fractions containing the majority of the 10 kDa mono-PEGylated IFN variants were combined (typically 7 mL) and then concentrated using a Vivaspin concentrator to 1 mL. Determination of the protein concentration for the three isolated PEG–IFN conjugates was accomplished using microBCA protein assay to calculate the final yields (Table 3).

Analysis by SDS–PAGE (silver stain detection) revealed there was no free IFN and RP-HPLC analysis indicated that the isolated conjugates were produced in high purity (Fig. S3† and 6C). No de-conjugation of PEG or free IFN was observed over a 28 day period for conjugates stored at 4 °C in 50 mM PBS, pH 7.4 (Fig. 6D, western blot analysis). Since SDS–PAGE analysis was consistent throughout,[32] no MALDI analysis of the conjugates were conducted.

The fact that the mono-PEGylated IFN variant was by far the major PEGylated species produced during each conjugation reaction meant a simple one-step purification process based on cation exchange chromatography was sufficient to purify the desired PEG–IFN conjugate. The simple purification process gave essentially the pure mono-PEGylated IFN variant (>97% by RP-HPLC) in relatively high yields (based on conversion) ranging from 21–33%. This is an excellent recovery for the observed conjugation conversions (ranging from 28–39%).

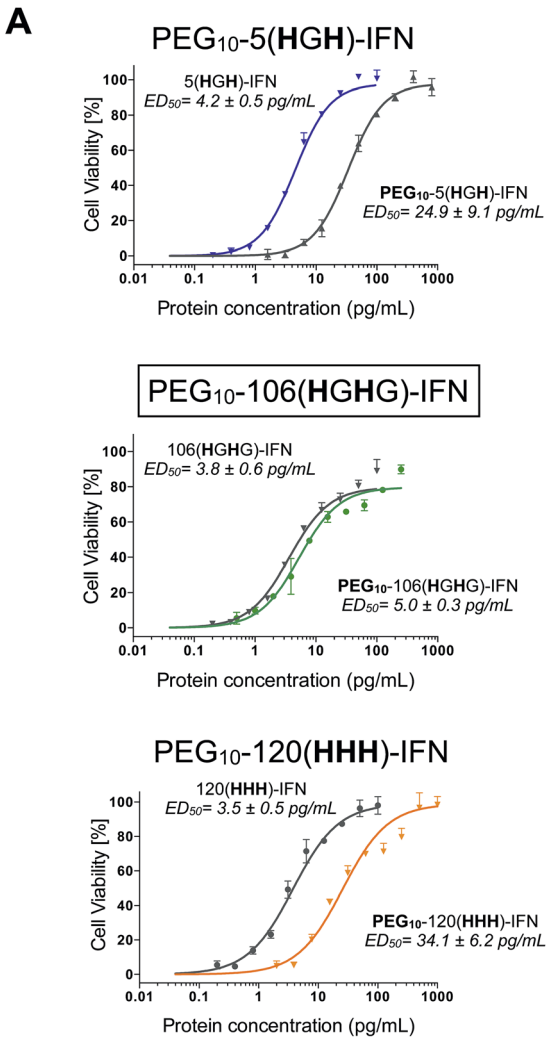
All three PEG–IFN conjugates are biologically active (Fig. 7A and B) in the A549/EMCV antiviral assay with specific activities calculated against the NIBSC standard. The site of PEG conjugation influences biological activity. As IFN binds to a heterodimeric cell surface receptor, the conjugation site is important to minimise competitive steric shielding effects of the PEG with IFN binding to its receptor. The conjugates displayed activity in the following order: PEG_{10}-106(**HGHG**)-IFN ≫ PEG_{10}-5(**HGH**)-IFN > PEG_{10}-120(**HHH**)-IFN.

The decreased activity observed with PEG_{10}-120(**HHH**)-IFN (10%) is consistent with the activity observed for the PEG positional isomers in PEG-Intron® and Pegasys® when PEG was conjugated to the D α-helix at position 121. The activity was 9% for PEG-Intron® and 2.2% for Pegasys®.[88,89] The greater reduction of bioactivity observed for Pegasys® may be due to the conjugation having been done with a branched PEG reagent (2 × 20 kDa PEG). The two separate PEG molecules, though hinged, appear to cause greater steric shielding of the receptor interaction site on IFN than a single linear PEG molecule.

The 17% activity of the PEG_{10}-5(**HGH**)-IFN variant is similar to that of the N-terminal $\mathbf{H_8}$-IFN that had previously been PEGylated.[64] Both PEG–IFN conjugates have the PEG near or at the N-terminus of the IFN, which further supports the

Table 3 Final isolated yields for the site-selectively PEGylated IFN variants

Conjugate	Final yield [%]
PEG_{10}-5(HGH)-IFN	25%
PEG_{10}-106(HGHG)-IFN	33%
PEG_{10}-120(HHH)-IFN	21%



B

Sample	n	ED50 (pg/mL) Averaged	SD	%CV	Specific activity to NIBSC IFNα-2a (MIU/mg) Averaged	SD	%CV	Relative activity [%]
PEG_{10}-5(HGH)-IFN	4	**24.9**	**9.1**	**36.7**	**37.9**	**6.7**	**17.7**	**17%**
5(**HGH**)-IFN	4	4.2	0.5	11.6	218.5	22.2	10.2	
PEG_{10}-106(HGH)-IFN	4	**5.0**	**0.3**	**6.8**	**139.6**	**5.5**	**3.9**	**74%**
106(**HGH**)-IFN	4	3.8	0.6	15.6	186.7	38.5	20.6	
PEG_{10}-120(HHH)-IFN	4	**34.2**	**6.2**	**18.2**	**24.9**	**3.8**	**15.2**	**10%**
120(**HHH**)-IFN	4	3.5	0.5	15.0	243.4	51.4	21.1	

Fig. 7 (A) Antiviral activity showed that all three of the internally histidine labelled PEG_{10}-IFN variants were active, and that the level of retained activity significantly depends on the site of conjugation. (B) Antiviral activity of the site-specifically PEGylated IFN variants revealed that the site of PEG conjugation has significant effect on the level of biological activity retained.

importance of the conjugation site on biological activity. Similar activity of 10–16% was observed with a IFN variant with cysteine inserted at position 5 for conjugates derived from PEG reagents with molecular weights ranging from 10–40 kDa.[15]

The activity of the PEG_{10}-106(**HGH**G)-IFN variant was similar to that of the un-PEGylated 106(**HGH**G)-IFN parent protein (5.0 ± 0.3 pg mL^{-1} *vs.* 3.8 ± 0.6 pg mL^{-1}). PEG_{10}-106(**HGH**G)-IFN displayed an exceptionally high activity of 74% relative to the unPEGylated protein. The conjugation site in the 106(**HGH**G)-IFN variant is located away from known receptor binding sites in a flexible and solvent accessible loop, which makes it possible for retaining most of the bioactivity after PEGylation. This level of antiviral activity for a stable interferon PEG conjugate is unprecedented. The PEG_{10}-106(**HGH**G)-IFN example demonstrates that a His_2-tag can serve as a selective site for bis-alkylation conjugation.

Conclusions

Interferon α2-a variants were engineered with histidine conjugation tags. In most cases two histidines were separated by a glycine (*e.g.* **HGH**) and this internal His-tag was placed at both the N-terminus and at different locations within the protein. We have termed these internal His-tags 'PEG-tags'. Bis-alkylation PEGylation was used to give conjugates that displayed *in vitro* antiviral activity that was dependent on the site of the PEG-tag. PEGylation was conducted with bis-alkylation reagent **3** derived from PEG with a molecular weight of 10 kDa.

Since histidine is a natural amino acid, it was possible to use routine methods of site-directed mutagenesis to make the IFN variants which were expressed in soluble form to give similar titres obtained for native IFN. PEGylation conversions ranged from 28–39% and a single step purification process gave essentially the pure PEG–IFN variant (>97% by RP-HPLC) in high recovery with isolated yields ranging from 21–33%.

The level of retained bioactivity was strongly dependent on the site of PEG conjugation. The highest biological activity of 74% was retained for the PEG_{10}-106(**HGH**G)-IFN which is unprecedented for a PEGylated IFN. The His_2-tag was placed at a location in IFN where it was expected there would be minimal interference of PEG during binding with the interferon receptors. The biological activity for the PEG_{10}-5(**HGH**)-IFN variant was 17% which is comparable to other PEGylated IFN conjugates at or near the N-terminus that have been described in the literature. The lowest retained activity (10%) was for PEG_{10}-120(**HHH**)-IFN which was prepared as a negative control as this IFN variant was thought to be involved in receptor binding.

The presence of two histidines in PEG-tags to generate a target for bis-alkylating PEGylation is a feasible approach for site-selective PEGylation. The use of a PEG-tag strategically placed in a protein can result in maximising the retention of the biological activity after protein modification.

Experimental

Preparation of PEG_{10}-mono-sulfone 3

The PEG mono-sulfone **3** derived from 10 kDa PEG was used in these studies. The PEG bis-sulfone **1** was first prepared[115] and was then incubated in 50 mM sodium phosphate buffer, pH 7.4, 2 mM EDTA, 150 mM NaCl at concentration of 10 mg mL^{-1} for a period of 6 h at 37 °C to provide the PEG mono-sulfone **3**.

Site-selective PEGylation of IFN

Conjugation reactions were conducted on 1 mg scale of each IFN variant. The IFN proteins: (i) 5(**HGH**)-IFN (1.5 mg mL^{-1}; 2 mL in 20 mM Tris, pH 8.0), (ii) 106(**HGH**G)-IFN (1.7 mg mL^{-1}; 2 mL in 20 mM Tris, pH 8.0) and (iii) 120(**HHH**)-IFN (1.6 mg mL^{-1}; 2 mL in 20 mM Tris, pH 8.0) were buffer exchange to an acetate-based PEGylation buffer (50 mM sodium acetate, pH 5.0 supplemented with 35 μM hydroquinone). Buffer exchange was conducted using a PD-10 desalting column (load: 2.5 mL) and eluted in 3.5 mL. The protein concentration determined by absorbance measurement at 280 nm were around 0.9 mg mL^{-1} for all samples (0.9 mg mL^{-1} for 5(**HGH**)-IFN; 0.92 mg mL^{-1} for 106(**HGH**G)-IFN and 0.9 mg mL^{-1} for 120(**HHH**)-IFN: 0.9 mg mL^{-1}). The PEG mono-sulfone **3** prepared at 10 mg mL^{-1} was then added to each reaction mixture using 5 molar equivalents: 226 μL for 5(**HGH**)-IFN 231 μL for 106(**HGH**G)-IFN; and 225 μL for 120(**HHH**)-IFN. The conjugation reaction was allowed to incubate for 16 h at 20 °C and then the reaction mixtures were treated with 25 mM sodium triacetoxyborohydride (19 mg) which was added to the reaction mixture as a solid. The reaction mixture was allowed to incubate for 45 min on ice. This sequence of adding sodium triacetoxyborohydride and 45 min incubation was repeated twice.

The reaction mixture was then buffer exchanged into 50 mM sodium acetate, pH 4.0, using a pre-equilibrated PD-10 desalting column as previously described[64] by loading 2.5 mL and eluting with 3.5 mL. The PD-10 column was again equilibrated with fresh buffer prior to the remaining 1.1 mL of reaction mixture being loaded, followed by addition of 1.4 mL of 50 mM sodium acetate, pH 4.0. The sample was then eluted with 3.5 mL of 50 mM sodium acetate, pH 4.0. The total volume of 7.5 mL of buffer exchanged reaction mixture was collected and subjected to ion exchange purification.

Cation exchange purification was performed on a MacroTrap SP HP (5 mL) column operated on an ÄKTAprime™ system. The column was firstly equilibrated with 30 mL of 50 mM sodium acetate, pH 4.0 (buffer A), followed by load of the PEGylation reaction mixture (7.5 mL). The flow-through was collected and the column was washed with 15 mL of buffer A to remove any residual PEG reagent. Subsequently, the column was washed with an increasing concentration of NaCl by applying a gradient elution of 50 mM sodium acetate, pH 4.0, 1 M NaCl (buffer B) from 0% to 100% typically over 30 min at 1 mL min^{-1}. Fractions containing the desired product of the mono-PEGylated IFN conjugate from each variant were combined and concentrated to 1.0 mL using a Vivaspin centrifugal concentrator (MWCO 10 000, centrifuged at 3000*g* at 4 °C).

Conflicts of interest

JWC is an employee of Abzena. KP, EL and RT no longer are employees of Abzena. SB was a co-founder of PolyTherics, a subsidiary of Abzena, but is not affiliated with either company.

Acknowledgements

We are grateful to Abzena for funding KP's PhD studies and for supporting the work described in this manuscript.

Notes and references

1 K. Wilding, A. Smith, J. Wilkerson, D. Bush, T. Knotts and B. Bundy, *ACS Synth. Biol.*, 2018, **7**, 510–521.
2 B. Pandey, M. Smith and J. Price, *Biomacromolecules*, 2014, **15**, 4643–4647.
3 B. Q. Shen, K. Xu, L. Liu, H. Raab, S. Bhakta, M. Kenrick, K. L. Parsons-Reponte, J. Tien, S. Yu, E. Mai, D. Li, J. Tibbitts, J. Baudys, O. M. Saad, S. J. Scales, P. J. McDonald, P. E. Hass, C. Eigenbrot, T. Nguyen, W. A. Solis, R. N. Fuji, K. M. Flagella, D. Patel, S. D. Spencer, L. A. Khawli, A. Ebens, W. L. Wong, R. Vandlen, S. Kaur, S. X. Sliwkowski, R. H. Scheller, P. Polakis and J. R. Junutula, *Nat. Biotechnol.*, 2012, 1–7.
4 M. Francis and I. Carrico, *Curr. Opin. Chem. Biol.*, 2010, **14**, 771–773.
5 G. Pasut and F. M. Veronese, *J. Controlled Release*, 2012, **161**, 461–472.
6 P. Turecek, M. Bossard, F. Schoetens and I. Ivens, *J. Pharm. Sci.*, 2016, **105**, 460–475.
7 H. Sato, *Adv. Drug Delivery Rev.*, 2002, **54**, 487–504.
8 D. Rabuka, *Curr. Opin. Chem. Biol.*, 2010, **14**, 790–796.
9 M. A. Gauthier and H. A. Klok, *Biomacromolecules*, 2010, **12**, 482–493.
10 J. Thom, D. Anderson, J. McGregor and G. Cotton, *Bioconjugate Chem.*, 2011, **22**, 1017–1020.
11 Q. Hu, F. Berti and R. Adamo, *Chem. Soc. Rev.*, 2016, **45**, 1691–1719.
12 E. Milczek, *Chem. Rev.*, 2017, **118**, 119–141.
13 Y. Wang and C. Wu, *Biomacromolecules*, 2018, **19**, 1804–1825.
14 M. S. Rosendahl, D. H. Doherty, S. J. Smith, S. J. Carlson, E. A. Chlipala and G. N. Cox, *Bioconjugate Chem.*, 2005, **16**, 200–207.
15 S. J. Bell, C. M. Fam, E. A. Chlipala, S. J. Carlson, J. I. Lee, M. S. Rosendahl, D. H. Doherty and G. N. Cox, *Bioconjugate Chem.*, 2008, **19**, 299–305.
16 R. Ohri, S. Bhakta, A. O'Donohue, J. Cruz-Chuh, S. Tsai, R. Cook, B. Wei, C. Ng, A. Wong, A. Bos, F. Farahi, J. Bhakta, T. Pillow, H. Raab, R. Vandlen, P. Polakis, Y. Liu, H. Erickson, J. Junutula and K. Kozak, *Bioconjugate Chem.*, 2018, **20**, 473–485.
17 M. Trivedi, J. Laurence and T. Siahaan, *Curr. Protein Pept. Sci.*, 2009, **10**, 614–625.
18 L. Zhang, C. Chou and M. Moo-Young, *Biotechnol. Adv.*, 2011, **29**, 923–929.
19 J. Junutula, S. Bhakta, H. Raab, K. E. Ervin, C. Eigenbrot, R. Vandlen, R. Scheller and H. B. Lowman, *J. Immunol. Methods*, 2008, **332**, 41–52.
20 P. England, *Biochemistry*, 2004, **43**, 11623–11629.
21 R. Connor and D. Tirrell, *J. Macromol. Sci., Polym. Rev.*, 2007, **47**, 9–28.
22 E. Sletten and C. Bertozzi, *Angew. Chem., Int. Ed.*, 2009, **48**, 6974–6998.
23 W. Zhang, G. Otting and C. Jackson, *Curr. Opin. Struct. Biol.*, 2013, **23**, 581–587.
24 O. Boutureira and G. J. L. Bernardes, *Chem. Rev.*, 2015, **115**, 2174–2195.
25 D. Young and P. Schultz, *ACS Chem. Biol.*, 2018, **13**, 854–870.
26 S. Schinn, W. Bradley, A. Groesbeck, J. Wu, A. Broadbent and B. Bundy, *Biotechnol. Bioeng.*, 2017, **114**, 2412–2417.
27 T. Kalstrup and R. Blunck, *Sci. Rep.*, 2015, **5**, 11866.
28 H. Cho, T. Daniel, Y. Buechler, D. Litzinger, Z. Maio, A. Putnam, V. Kraynov, B. Sim, S. Bussell and T. Javahishvili, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 9060–9065.
29 Q. Wang and L. Wang, *J. Am. Chem. Soc.*, 2008, **130**, 6066–6067.
30 R. Martin, S. BJ, Y. Kwon, J. Kay, R. Davis, P. Thomas, N. Majewska, C. Chen, R. Marcum, M. Weiss, A. Stoddart, M. Amiram, A. Charna, J. Patel, F. Isaacs, N. Kelleher, S. Hong and M. Jewett, *Nat. Commun.*, 2018, **9**, 1203.
31 G. T. Debelouchina and T. W. Muir, *Q. Rev. Biophys.*, 2017, **50**, 1–41.
32 N. Nischan and C. Hackenberger, *J. Org. Chem.*, 2014, **79**, 10727–10733.
33 Z. Harris and J. Gitlin, *Am. J. Clin. Nutr.*, 1996, **63**, 836S–841S.
34 M. Smith, M. Easton, P. Everett, G. Lewis, M. Payne, V. RiverosMoreno and G. Allen, *Int. J. Pept. Protein Res.*, 1996, **48**, 48–55.
35 C. Pickens, S. Johnson, M. Pressnall, M. Leon and C. Berkland, *Bioconjugate Chem.*, 2018, **29**, 686–701.
36 A. Wang, N. Nairn, M. Marielli and K. Grabstein, in *Protein Engineering*, ed. P. Kaumaya, InTech, Rijeka, Croatia, 2012.
37 K. Hamblett, P. Senter, D. Chace, M. Sun, J. Jenox, C. Cerveny, K. Kissler, S. Bernhardt, A. Kopcha, R. F. Zabinski, D. Meyer and J. A. Francisco, *Clin. Cancer Res.*, 2004, **10**, 7063–7070.
38 A. Beck, L. Goetsch, C. Dumontet and N. Corvaïa, *Nat. Rev. Drug Discovery*, 2017, **16**, 315–337.
39 E. Sletten and C. Bertozzi, *Org. Lett.*, 2008, **10**, 3097–3099.
40 E. Kozma, I. Nikic, B. Varga, I. Aramburu, J. Kang, O. Fackler, E. Lemke and P. Kele, *ChemBioChem*, 2016, **17**, 1518–1524.
41 T. Satomaa, H. Pynnonen, A. Vilkman, T. Kotiranta, V. Pitkanen, A. Heiskanen, B. Herpers, L. Price, J. Helin and J. Saarinen, *Antibodies*, 2018, **7**, 15.
42 R. Lyon, T. Bovee, S. O. Doronina, P. Burke, J. Hunter, H. Neff-LaFord, M. Jonas, M. Anderson, J. Setter and P. Senter, *Nat. Biotechnol.*, 2015, **33**, 733–735.
43 J. J. Kopchick, C. Parkinson, E. C. Stevens and P. J. Trainer, *Endocr. Rev.*, 2002, **23**, 623–646.
44 Y. Yamamoto, Y. Tsutsumi, Y. Yoshioka, T. Nishibata, K. Kobayashi, T. Okamoto, Y. Mukai, T. Shimizu, S. Nakagawa, S. Nagata and T. Mayumi, *Nat. Biotechnol.*, 2003, **21**, 546–552.
45 F. Metzger, W. Sajid, S. Saenger, C. Staudenmaier, C. Poel, B. Sobottka, A. Schuler, M. Sawitzky, R. Poirier, D. Tuerck,

E. Schick, A. Schaubmar, F. Hesse, K. Amrein, H. Loetscher, G. Lynch, A. Hoeflich, P. DeMeyts and H. Schoenfeld, *J. Biol. Chem.*, 2011, **286**, 19501–19510.

46 T. Morishige, Y. Yoshioka, H. Inakura, A. Tanabe, X. Yao, S. Tsunoda, Y. Tsutsumi, Y. Mukai, N. Okada and S. Nakagawa, *Biochem. Biophys. Res. Commun.*, 2010, **393**, 888–893.

47 A. deGraaf, M. Kooijman, W. Hennink and E. Mastrobattista, *Bioconjugate Chem.*, 2009, **20**, 1281–1295.

48 B. Wiman and P. Wallen, *Thromb. Res.*, 1977, **10**, 213–222.

49 G. Gitlin, E. Bayer and M. Wilchek, *Biochem. J.*, 1987, **242**, 923–926.

50 F. Rozen, J. Pelletier, H. Trachsel and N. Sonenberg, *Mol. Cell. Biol.*, 1989, **9**, 4061–4063.

51 M. Matos, B. Oliveira, N. Martínez-Sáez, A. Guerreiro, P. Cal, J. Bertoldo, M. Maneiro, E. Perkins, J. Howard, M. Deery, J. Chalker, F. Corzana, G. Jiménez-Osés and G. Bernardes, *J. Am. Chem. Soc.*, 2018, **140**, 4004–4017.

52 S. Shaunak, A. Godwin, J. W. Choi, S. Balan, E. Pedone, D. Vijayarangam, S. Heidelberger, I. Teo, M. Zloh and S. Brocchini, *Nat. Chem. Biol.*, 2006, 312–313.

53 S. Balan, J. W. Choi, A. Godwin, I. Teo, C. M. Laborde, S. Heidelberger, M. Zloh, S. Shaunak and S. Brocchini, *Bioconjugate Chem.*, 2007, **18**, 61–76.

54 M. E. B. Smith, F. F. Schumacher, C. P. Ryan, L. M. Tedaldi, D. Papaioannou, G. Waksman, S. Caddick and J. R. Baker, *J. Am. Chem. Soc.*, 2010, **132**, 1960–1965.

55 Z. Li, R. Huang, H. Xu, J. Chen, Y. Zhan, X. Zhou, H. Chen and B. Jiang, *Org. Lett.*, 2017, **19**, 4972–4975.

56 S. Kuan, T. Wang and T. Weil, *Chem.–Eur. J.*, 2016, **22**, 17112–17129.

57 C. Bahou, D. Richards, A. Maruani, E. Love, F. Javaid, S. Caddick, J. Baker and V. Chudasama, *Org. Biomol. Chem.*, 2018, **16**, 1359–1366.

58 G. Badescu, P. Bryant, M. Bird, K. Henseleit, J. Swierkosz, V. Parekh, R. Tommasi, E. Pawlisz, K. Jurlewicz, M. Farys, N. Camper, X. Sheng, M. Fisher, R. Grygorash, A. Kyle, A. Abhilash, M. Frigerio, J. Edwards and A. Godwin, *Bioconjugate Chem.*, 2014, **25**, 1124–1136.

59 P. Bryant, M. Pabst, G. O. Badescu, M. Bird, W. McDowell, E. Jamieson, J. Swierkosz, K. Jurlewicz, R. Tommasi, K. Henseleit, X. Sheng, N. Camper, A. Manin, K. Kozakowska, K. Peciak, E. Laurine, R. Gurygorash, A. Kyle, D. Morris, V. Parekh, A. Abhilash, J. Choi, J. Edwards, M. Frigerio, M. Baker and A. Godwin, *Mol. Pharmaceutics*, 2015, **12**, 1872–1879.

60 M. Pabst, W. McDowell, A. Manin, A. Kyle, N. Camper, E. DeJuan, V. Parekh, F. Rudge, H. Makwana, T. Kantner, H. Parekh, A. Michelet, X. Sheng, G. Popa, C. E. Tucker, F. Khayrzad, D. Pollard, K. Kozakowska, R. Resende, A. Jenkins, F. Simoes, D. Morris, P. Williams, G. Badescu, M. Baker, M. Bird, M. Frigerio and A. Godwin, *J. Controlled Release*, 2017, **253**, 160–164.

61 H. Khalili, P. Khaw, R. Lever, A. Godwin and S. Brocchini, *Bioconjugate Chem.*, 2013, **24**, 1870–1882.

62 H. Khalili, P. Khaw and S. Brocchini, *Biomater. Sci.*, 2016, **4**, 943–947.

63 H. Khalili, R. Lee, P. Khaw, S. Brocchini and A. D. Dick, *Sci. Rep.*, 2016, **6**, 36905.

64 Y. Cong, E. Pawlisz, P. Bryant, S. Balan, E. Laurine, R. Tommasi, R. Singh, S. Dubey, K. Peciak, M. Bird, A. Sivasankar, J. Swierkosz, M. Muroni, S. Heidelberger, M. Farys, F. Khayrzad, J. Edwards, G. Badescu, I. Hodgson, C. Heise, S. Somavarapu, J. Liddell, K. Powell, M. Zloh, J.-w. Choi, A. Godwin and S. Brocchini, *Bioconjugate Chem.*, 2012, **23**, 248–263.

65 E. A. Woestenenk, M. Hammarström, S. Van den Berg, T. Härd and H. Berglund, *J. Struct. Funct. Genomics*, 2004, **5**, 217–229.

66 M. Li and D. Huang, *Protein Expression Purif.*, 2007, **53**, 119–123.

67 A. Tinazli, J. Tang, R. Caliokas, S. Picuric, S. Lata, J. Piehler, B. Liedberg and R. Tampé, *Chem.–Eur. J.*, 2005, **11**, 5249–5259.

68 C. Hauser and R. Tsien, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 3693–3697.

69 Y. Lai, Y. Chang, L. Hu, A. Chao, Z. Du, J. Tanner, M. Cheye, C. Qian, K. Ng, H. Li and H. Sun, *Proc. Natl. Acad. Sci. U. S. A.*, 2015, **112**, 2948–2953.

70 G. Meredith, H. Wu and N. L. Allbritton, *Bioconjugate Chem.*, 2004, **15**, 969–982.

71 S. Uchinomiya, H. Nonaka, S. Fujishima, S. Tsukiji, A. Ojida and I. Hamachi, *Chem. Commun.*, 2009, 5880–5882.

72 V. Bolanos-Garcia and O. Davies, *Biochim. Biophys. Acta, Gen. Subj.*, 2006, **1760**, 1304–1313.

73 L. Kozlowski, *Nucleic Acids Res.*, 2017, **45**, D1112–D1116.

74 B. Zhang, H. Xu, J. Chen, Y. Zheng, Y. Wu, L. Si, L. Wu, C. Zhang, G. Xia, L. Zhang and D. Zhou, *Acta Biomater.*, 2015, **19**, 100–111.

75 A. Isaacs and J. Lindenmann, *Proc. R. Soc. B*, 1957, **147**, 258–267.

76 J. Vilcek and M. Feldmann, *Trends Pharmacol. Sci.*, 2004, **25**, 201–209.

77 V. Tayal and B. Kalra, *Eur. J. Pharmacol.*, 2008, **579**, 1–12.

78 S. Presnell and F. Cohen, *Proc. Natl. Acad. Sci. U. S. A.*, 1989, **86**, 6592–6596.

79 S. Kamtekar and M. Hecht, *FASEB J.*, 1995, **9**, 1013–1022.

80 W. Klaus, B. Gsell, A. Labhardt, B. Wipf and H. Seen, *J. Mol. Biol.*, 1997, **274**, 661–675.

81 S. Polyak, G. Forsberg, B. Forbes, K. McNeil, S. Aplin and J. Wallace, *Protein Eng.*, 1997, **10**, 615–619.

82 K. Peciak, R. Tommasi, J. Choi, S. Brocchini and E. Laurine, *Protein Expression Purif.*, 2014, **99**, 18–26.

83 H. Khalili, A. Godwin, J. Choi, R. Lever and S. Brocchini, *Bioconjugate Chem.*, 2012, **23**, 2262–2277.

84 O. Kinstler, G. Molineux, M. Treuheit, D. Ladd and C. Gegg, *Adv. Drug Delivery Rev.*, 2002, **54**, 477–485.

85 M. Cindric, T. Cepo, N. Galic, M. Bukvic-Krajacic, N. Tomczyk, J. P. C. Vissers, L. Bindila and J. Peter-Katalinic, *J. Pharm. Biomed. Anal.*, 2007, **44**, 388–395.

86 J. Wang, T. Hu, Y. Liu, G. Zhang, G. Ma and Z. Su, *Anal. Biochem.*, 2011, **412**, 114–116.

87 C. Rosen and M. Francis, *Nat. Chem. Biol.*, 2017, **13**, 697–705.

88 Y. Wang, S. Youngster, J. Bausch, R. Zhang, C. McNemar and D. Wyss, *Biochemistry*, 2000, **39**, 10634–10640.
89 S. Foser, A. Schacher, K. Weyer, D. Brugger, E. Dietel, S. Marti and T. Schreitmuller, *Protein Expression Purif.*, 2003, **30**, 78–87.
90 *The interferons: characterization and application*, ed. A. Meager, Wiley-VCH Verlag, Weinheim, 2006.
91 S. Quadt-Akabayov, J. Chill, Y. Levy, N. Kessler and J. Anglister, *Protein Sci.*, 2006, **15**, 2656–2668.
92 J. Bazan, *Proc. Natl. Acad. Sci. U. S. A.*, 1990, **87**, 18.
93 S. Pestka, *Biopolymers*, 2000, **55**, 254–287.
94 J. Chill, S. Quadt and J. Anglister, *Biochemistry*, 2004, **43**, 10127–10137.
95 C. Krause and S. Pestka, *Pharmacol. Ther.*, 2005, **106**, 299–346.
96 Y. S. Wang, S. Youngster, M. Grace, J. Bausch and D. F. Wyss, *Adv. Drug Delivery Rev.*, 2002, **54**, 547–570.
97 M. Grace, S. Youngster, G. Gitlin, W. Sydor, L. Xie, L. Westreich, S. Jacobs, D. Brassard, J. Bausch and R. Bordens, *J. Interferon Cytokine Res.*, 2001, **21**, 1103–1115.
98 M. J. Roberts, M. D. Bentley and J. M. Harris, *Adv. Drug Delivery Rev.*, 2002, **54**, 459–476.
99 P. Bailon, A. Palleroni, C. A. Schaffer, C. L. Spence, W. Fung, J. E. Porter, G. K. Ehrlich, W. Pan, Z. Xu, M. W. Modi, A. Farid and W. Berthold, *Bioconjugate Chem.*, 2001, **12**, 195–202.
100 A. Basu, K. Yang, M. Wang, S. Liu, R. Chintala, T. Palm, H. Zhao, P. Peng, D. Wu, Z. Zhang, J. Hua, M. Hsieh, Z. Zhang, G. Petti, X. Y. Li, A. Janjua, M. Mendez, J. Liu, C. Longley, Z. Zhang, M. Mehlig, V. Borowski, M. Viswanathan and D. Filpula, *Bioconjugate Chem.*, 2006, **17**, 618–630.
101 C. Dhalluin, A. Ross, L. Leuthold, S. Foser, B. Gsell, F. Muller and H. Senn, *Bioconjugate Chem.*, 2005, **16**, 504–517.
102 C. Dhalluin, A. Ross, W. Huber, P. Gerber, D. Brugger, B. Gsell and H. Senn, *Bioconjugate Chem.*, 2005, **16**, 518–527.
103 G. Adolf, I. Kalsner, H. Ahorn, I. Maurer-Fogy and K. Cantell, *Biochem. J.*, 1991, **276**, 511–518.
104 T. Nyman, N. Kalkkinen, H. Tolo and J. Helin, *Eur. J. Biochem.*, 1998, **253**, 485–493.
105 S. Akabaov, Z. Biron, P. Lamken, J. Piehler and J. Anglister, *Biochemistry*, 2010, **49**, 687–695.
106 I. Nedelman, S. Akabayov, E. Schnur, Z. Biron, R. Levy, Y. Xu, D. Yang and J. Anglister, *Biochemistry*, 2010, **49**, 5117–5133.
107 G. Foster, *Aliment. Pharmacol. Ther.*, 2004, **20**, 825–830.
108 G. Badescu, P. Bryant, J. Swierkosz, F. Khayrzad, E. Pawlisz, M. Farys, Y. Cong, M. Muroni, N. Rumpf, S. Brocchini and A. Godwin, *Bioconjugate Chem.*, 2014, **25**, 460–469.
109 H. Lee, H. Jang, S. Ryu and T. Park, *Pharm. Res.*, 2003, **20**, 818–825.
110 J. Tong, K. Zhong, H. Tian, J. Gao, X. Xu, X. Yin and W. Yao, *Int. J. Biol. Macromol.*, 2010, **46**, 331–336.
111 I. A. Puchkov, N. V. Kononova, A. I. Bobruskin, D. I. Bairamashvili, V. A. Martyanov and A. M. Shuster, *Russ. J. Bioorg. Chem.*, 2012, **38**, 479–487.
112 R. P. Lyon, J. R. Setter, T. D. Bovee, S. O. Doronina, J. H. Hunter, M. E. Anderson, C. L. Balasubramanian, S. M. Duniho, C. I. Leiske, F. Li and P. D. Senter, *Nat. Biotechnol.*, 2014, 1–7, DOI: 10.1038/nbt.2968.
113 M. Morais, J. Nunes, K. Karu, N. Forte, I. Benni, M. Smith, S. Caddick, V. Chudasama and D. Baker, *Org. Biomol. Chem.*, 2017, **15**, 2947–2952.
114 D. Kalia, S. Pawar and J. Thopate, *Angew. Chem., Int. Ed.*, 2017, **56**, 1885–1889.
115 S. Brocchini, S. Balan, A. Godwin, J. W. Choi, M. Zloh and S. Shaunaik, *Nat. Protoc.*, 2006, **1**, 2241–2252.