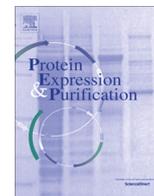




Contents lists available at ScienceDirect

## Protein Expression and Purification

journal homepage: [www.elsevier.com/locate/yprep](http://www.elsevier.com/locate/yprep)

## Functional Siglec lectin domains from soluble expression in the cytoplasm of *Escherichia coli*



Johannes M. Pröpster<sup>a</sup>, Fan Yang<sup>b,c</sup>, Beat Ernst<sup>b</sup>, Frédéric H.-T. Allain<sup>a,\*</sup>, Mario Schubert<sup>a,d,\*</sup>

<sup>a</sup>Institute of Molecular Biology and Biophysics, ETH Zürich, CH-8093 Zürich, Switzerland

<sup>b</sup>Institute of Molecular Pharmacy, Pharmazentrum, University of Basel, CH-4056 Basel, Switzerland

<sup>c</sup>Department of Molecular & Medical Pharmacology, David Geffen School of Medicine at UCLA, CA 90095, USA<sup>1</sup>

<sup>d</sup>Department of Molecular Biology, University of Salzburg, A-5020 Salzburg, Austria<sup>1</sup>

### ARTICLE INFO

#### Article history:

Received 8 December 2014

and in revised form 9 January 2015

Available online 24 January 2015

#### Keywords:

Siglec

I-type lectin

Recombinant expression

Disulfide bond

Carbohydrate recognition

Glycan-binding receptor

### ABSTRACT

Siglecs (sialic acid-binding immunoglobulin-like lectins) are a family of mammalian cell-surface receptors that are involved in cell–cell interactions and signaling functions, primarily expressed on cells of the immune system. Key to their function is their specific binding of distinct sialylated glycan ligands mediated via an N-terminal carbohydrate recognition (lectin) domain. Studies concerning the molecular basis of their individual carbohydrate specificities are rare due to the absence of suitable recombinant expression methods for producing these disulfide-containing proteins in sufficient quantities required for their in-depth *in vitro* characterization. We established an efficient *E. coli*-based expression and purification method for Siglec lectin domains, utilizing the *trxB gor* suppressor strain Rosetta-gami B (DE3) in which proper folding with intact disulfide bonds was achieved in the cytoplasm. The approach is demonstrated for human Siglec-7, -8 and -9 lectin domains and works equally well for expression in nutrient-rich (LB) or minimal growth medium, allowing stable-isotope labeling for NMR studies. The recombinant proteins were properly folded as proven by 2D <sup>1</sup>H–<sup>15</sup>N HSQC NMR spectroscopy and by thermal unfolding followed by CD spectroscopy, and functionally active as confirmed by monitoring ligand binding using NMR titration experiments. Our method enables efficient production of homogeneous and active protein samples in milligram quantities. Its implementation will significantly enhance future structure–function studies of this important class of immune-modulating receptors and will support a variety of applications including screening for natural and synthetic ligands or the development of fluorescently-labeled molecular tools for glycan ligand detection or flow-cytometric cell sorting.

© 2015 Elsevier Inc. All rights reserved.

### Introduction

Sialic acid-binding immunoglobulin-like lectins (Siglecs) are a prominent family of carbohydrate-binding cell-surface receptors predominantly found on immune cells [1,2]. Most Siglecs function as signaling co-receptors regulating the activity of cells responsible for innate and adaptive immune responses [3]. Their expression is often highly cell type-restricted, enabling their participation in discrete cellular functions, such as the down-regulation of B-cell activity [4] or the inhibition of natural killer (NK) cell cytotoxicity [5]. Owing to their capacity to interfere with different immune

cellular functions, Siglecs are gaining increasing attention as therapeutic targets for numerous immune and inflammatory disorders [6]. To date, a monoclonal antibody targeting Siglec-3 (CD33) is approved for the treatment of acute myeloid leukemia (AML) and several antibodies targeting Siglec-2 (CD22) for treating non-Hodgkin's lymphoma are in clinical trials [7]. In addition to antibody-based therapeutics, more recent strategies concentrate on targeting Siglecs with synthetic glycan-based ligands. Despite extensive evidence for the involvement of Siglecs in a variety of crucial immune regulatory events, their precise modes of actions are still not fully understood and require further investigations. Central to the function of Siglecs is their specific recognition of sialylated glycoconjugates by a single N-terminal lectin domain. Their carbohydrate targets can either be present on the same cell or even on Siglecs themselves (cis ligands) or exposed on neighboring cells, external molecules or even on the surface of pathogens (trans ligands). To analyze such complex interaction networks, a divide and conquer approach focusing on individual binding events

\* Corresponding authors at: Department of Molecular Biology, University of Salzburg, A-5020 Salzburg, Austria. Tel.: +43 662 80447243; fax: +43 662 80447209 (M. Schubert). Institute of Molecular Biology and Biophysics, ETH Zürich, CH-8093 Zürich, Switzerland. Tel.: +41 44 6333940; fax: +41 44 6331294 (F.H.-T. Allain).

E-mail addresses: [allain@mol.biol.ethz.ch](mailto:allain@mol.biol.ethz.ch) (F.H.-T. Allain), [mario.schubert@sbg.ac.at](mailto:mario.schubert@sbg.ac.at) (M. Schubert).

<sup>1</sup> Present address.

seems to be most promising. Insights into Siglec carbohydrate recognition at the atomic level are best obtained by studying the interaction of an isolated lectin domain that lacks glycans with well-defined oligosaccharides. For *in vitro* studies, sufficient quantities of homogeneous and functional proteins are required that can only be obtained from recombinant protein expression. However, the recombinant production of Siglec lectin domains is challenging, in particular because these domains contain native disulfide bonds. Several Siglecs have been recombinantly produced using different mammalian cell lines, thereby exploiting the mammalian folding machinery and secretory pathway to facilitate proper protein folding and disulfide bond formation [8–10]. However, such eukaryotic expression systems are typically prone to non-uniform glycosylation, resulting in multiple glycoforms of the target protein. An adaption of these systems for producing isotopically labeled proteins for NMR spectroscopic studies would be immoderately expensive. Economical yields of homogeneous, non-glycosylated and either isotopically- or selenomethionine-labeled proteins are best achieved by recombinant expression in prokaryotic host systems. The production of Siglec lectin domains in *Escherichia coli* proved to be difficult and only few attempts have been reported so far: the expression of human Siglec-5 and -7 lectin domains in inclusion bodies [11–13]. Their refolding required the development of individual, complicated protocols and yielded only small fractions of correctly folded protein. The use of alternative expression strategies, in particular those known to promote the formation of stable disulfide bonds, such as targeting the expression into the oxidative *E. coli* periplasm or specific cell-free expression systems, has not been reported for Siglecs so far. Finding an efficient way to express Siglec lectin domains in a soluble and functionally active state in bacteria would be a major breakthrough for studying the molecular basis of Siglec carbohydrate recognition and would pave the way for a range of biotechnological and pharmaceutical applications.

Here we present a simple and robust method for the recombinant production of soluble, folded and functionally active lectin domains of Siglecs in *E. coli*, using the *trxB gor* double mutant strain Rosetta-gami B (DE3) that allows disulfide bond formation in its oxidative cytoplasm. The efficiency of the technique was demonstrated with human Siglec-7, -8 and -9, yielding highly pure and homogeneous recombinant protein samples in milligram quantities. Thermal stabilities of the purified proteins were assessed by measuring temperature-induced unfolding transitions using far-UV CD spectroscopy, indicating stably folded proteins suitable for *in vitro* studies. Efficient isotope labeling was achieved and NMR spectroscopy was applied to monitor binding between the lectin domains and their oligosaccharide targets in solution, thereby proving that the recombinant proteins are functionally active. This work provides the fundamental basis for future structural and functional investigations of this important class of immunomodulatory cell-surface receptors.

## Materials and methods

### Cloning of expression plasmids

Full-length cDNA clones of Siglec-7 (IRATp970F1050D; NCBI GenBank entry: BC028150), Siglec-8 (IRAU969G05104D; NCBI GenBank entry: BC053319) and Siglec-9 (IRATp970B1247D; NCBI GenBank entry: BC035365; corresponding to natural variant *Lys100-Glu*), purchased from imaGenes GmbH, Germany were applied to site-directed mutagenesis to replace Cys41, Cys42 or Cys36 in Siglec-7, -8 or -9 respectively, by a serine. Resulting cDNA mutants served as templates for PCR amplification of inserts of the isolated lectin domains, encompassing the coding regions for

Gln19-His148, Met17-His155 and Gln18-His144 of the full-length protein sequences (or corresponding to Gln1-His130, Met1-His139 and Gln1-His127 of the mature protein sequences) of Siglec-7, -8 and -9, respectively. Used primer sets incorporate an NdeI restriction site upstream, and an XhoI restriction site downstream of the amplified fragment. The NdeI restriction site contains an ATG start codon, which was directly utilized to initiate the coding region for Siglec-8, starting from its natural N-terminal residue (Met17). Forward primers used for amplification of Siglec-7 and -9 constructs included an additional codon (GCC) after the NdeI restriction site, resulting in amplicons encoding for two additional, non-natural residues (MetGly) preceding the natural N-termini of the mature protein sequences. Reverse primers incorporate a two residue spacer (GlySer) followed by a thrombin cleavage site (Leu-ValProArg↓GlySer), directly downstream of the Siglec lectin domains and upstream of the XhoI restriction site. The PCR amplicons were digested with restriction enzymes NdeI and XhoI (New England Biolabs) and cloned between matching sites of expression vector pET-43.1(a) (Novagen), yielding final expression plasmids named pET43\_Sig7v1\_C41S, pET43\_Sig8v1\_C42S and pET43\_Sig9v1\_C36S. Constructs were confirmed by DNA sequence analysis (Microsynth AG, CH). All cloning steps were performed using standard molecular biology techniques [14]. *E. coli* strain TOP10 (Invitrogen) was used for cloning, amplification and maintenance of plasmids. The used cDNA clones, bacterial strains, plasmids and oligonucleotide primers are listed in [Supplementary Table S1](#).

### Analytical expressions to screen for soluble expression conditions

pET43\_Sig9v1\_C36S was transformed into BL21-Codon-Plus(DE3)-RIL or Rosetta-gami B(DE3), streaked onto LB (Luria-Bertani broth) agar plates containing 50 µg/ml of carbenicillin and 34 µg/ml of chloramphenicol, or for Rosetta-gami B(DE3) additionally containing 12.5 µg/ml of tetracycline and 15 µg/ml of kanamycin and incubated at 37 °C. A single, freshly grown colony was inoculated into 10 ml LB medium containing 50 µg/ml of carbenicillin and 34 µg/ml of chloramphenicol and grown at 37 °C and 250 rpm for 4.5 h. The 10 ml LB-culture was transferred into 100 ml of 37 °C-pre-warmed M9 minimal medium, supplemented with 1 g/L of NH<sub>4</sub>Cl, 4 g/L glucose, vitamins (Gibco MEM Vitamins, Life Technologies) and trace metals [15], 50 µg/ml of carbenicillin and 34 µg/ml of chloramphenicol and grown overnight at 37 °C and 250 rpm. The overnight-culture was used to inoculate 1.2 L of 37 °C-pre-warmed M9 minimal medium (same supplementary ingredients as described above) to a starting optical density at 600 nm (OD<sub>600</sub>) of 0.05 and grown at 30 °C and 150 rpm. At OD<sub>600</sub> of 0.65 the culture was split into 2 × 600 ml fractions and induced with 1 or 0.1 mM isopropyl-β-D-1-thiogalactoside (IPTG). Each of the 600 ml cultures was further split into 4 × 150 ml fractions and grown at different temperatures (37, 30, 22 and 15 °C). 10 ml samples were taken at different time-points (before induction and 2, 4, 8, 16, 20 and 38 h after induction), centrifuged at 4000 rpm for 10 min at 4 °C and stored at -20 °C. For analysis, cell pellets were thawed on ice, resuspended in 10 ml of ice-cold lysis-buffer (50 mM Tris-HCl, 500 mM NaCl, pH 8.0) and sonicated on ice (5 × 20 s pulses with 20 s pause between pulses). 2 ml of the lysates were centrifuged at 14,000 rpm for 1 h at 4 °C to remove insoluble material. Then, 1.5 ml of the soluble supernatant was transferred into 7.5 ml acetone (pre-chilled to -20 °C), shortly vortexed, incubated at -20 °C for 30 min for complete protein precipitation and centrifuged at 4000 rpm for 1 h at 4 °C. The supernatant was discarded and pellets were air-dried for 30 min at room temperature then resuspended in 37.5 µl of reducing 4× SDS sample buffer (0.15 M Tris-HCl pH 7.0, 16% v/v β-mercaptoethanol, 0.05% w/v bromophenol blue, 35% v/v glycerol, 4% SDS) plus 112.5 µl of phosphate-buffered-saline (PBS) pH 7.4. Samples were

subjected to SDS–PAGE (15%) followed by Coomassie blue-staining (Supplementary Fig. S1). Western blot analysis was performed following the protocol as described by Towbin et al. [16] using a primary monoclonal mouse anti-polyHistidine antibody (Sigma–Aldrich) and a secondary goat anti-mouse IgG antibody, conjugated to alkaline phosphatase (Sigma–Aldrich). Protein bands were visualized in alkaline phosphatase reaction buffer (100 mM Tris–HCl pH 9.5, 100 mM NaCl, 5 mM MgCl<sub>2</sub>) using NBT/BCIP (Roche, CH) as substrate for staining.

#### Preparative expression of soluble His-tagged Siglec lectin domains

pET43\_Sig7v1\_C41S, pET43\_Sig8v1\_C42S or pET43\_Sig9v1\_C36S were transformed into Rosetta-gami B(DE3), streaked onto LB agar plates containing 50 µg/ml of carbenicillin, 34 µg/ml of chloramphenicol, 12.5 µg/ml of tetracycline and 15 µg/ml of kanamycin and incubated at 37 °C for 25–30 h. All liquid culture media were supplemented with 50 µg/ml of carbenicillin and 34 µg/ml of chloramphenicol. For expression of unlabeled Siglec lectin domains a single, freshly grown colony was inoculated into LB medium and grown overnight at 37 °C and 250 rpm. The 37 °C-pre-warmed LB main cultures were inoculated to a starting OD<sub>600</sub> of 0.08 and grown at 37 °C and 130 rpm until an OD<sub>600</sub> of ≈0.7, and then the shaker temperature was switched to 15 °C. Expressions were typically performed in 1–6 L culture medium, with 1 L per 5-L Erlenmeyer flask. After reaching temperatures of ≤20 °C the expression was induced by addition of 1 mM IPTG, and cells were further grown at 15 °C for 40 h. Cells were harvested by centrifugation at 7000 rpm for 30 min at 4 °C. For expression of uniformly (<sup>15</sup>N) isotope-labeled proteins, a 10 ml LB medium starter culture inoculated with a single colony, was grown for 4–5 h at 37 °C and 250 rpm, and then diluted (1/10) into 37 °C-pre-warmed M9 minimal medium, supplemented with 1 g/L of <sup>15</sup>NH<sub>4</sub>Cl, 4 g/L glucose, vitamins (Gibco MEM Vitamins, Life Technologies) and trace metals [15], which was then grown overnight at 37 °C and 250 rpm. The pre-culture was used to inoculate the 37 °C-pre-warmed M9 minimal medium main cultures (same supplementary ingredients as described above), which were grown, induced and harvested with similar procedures as described for the unlabeled expression in LB medium, except that harvesting was performed 60 h after induction. Cell pellets were either directly used for purification or flash-frozen in liquid N<sub>2</sub> and stored at –80 °C.

#### Protein purification

Bacterial cell pellets were thawed on ice, resuspended in ice-cold buffer A (20 mM sodium phosphate, 500 mM NaCl, 20 mM imidazole, pH 7.4) additionally supplemented with Roche cOmplete EDTA-free protease-inhibitor mix (Roche, CH) and subsequently lysed by 3–4 passes through a pre-chilled microfluidizer (model: M-110 L; Microfluidics Inc., USA) operated at ≈14,000 psi chamber pressure. Insoluble material was removed by centrifugation at 17,000 rpm for 30 min at 4 °C. The soluble supernatant was passed through a 0.45-µm pore-size syringe filter (Filtropur S 0.45, Sarstedt) directly before loading on a prepacked Ni<sup>2+</sup>–NTA agarose column (HisTrap FF, 1 ml or 5 ml; GE Healthcare) pre-equilibrated with buffer A using an ÄKTA Prime FPLC system (GE Healthcare) at 4 °C. After washing with 20–30 column volumes of buffer A the protein was eluted with a linear gradient (0–100%) of buffer B (20 mM sodium phosphate, 500 mM NaCl, 500 mM imidazole, pH 7.4) over a length of 20 column volumes. Peak fractions containing the target protein were pooled (Supplementary Fig. S2) and concentrated by centrifugation at 4 °C in a Vivaspin 6 concentrator (5 kDa MWCO; Sartorius) to a volume of 220 µl per 1 L of the original cell culture volume. The C-terminal His<sub>6</sub>-tag was removed by thrombin cleavage without a preceding

buffer exchange, by adding thrombin (T4648; Sigma–Aldrich), dissolved at 1 NIH unit/µl in buffer A, at a dose of 10–15 NIH units of enzyme per 220 µl of concentrated eluate for overnight digestion at 4 °C. The cleaved protein solution was then directly subjected to a Superdex 75 10/300 GL (GE Healthcare) gel filtration column connected to an ÄKTA Purifier FPLC system (GE Healthcare) which was sufficient for separation of the cleaved affinity tag and other remaining impurity traces and to transfer the protein into the final buffer containing 20 mM potassium phosphate, 40 mM NaCl, pH 7.0 (Supplementary Fig. S2). Identities of the purified proteins and the presence of their intra-domain disulfide bond were confirmed by SDS–PAGE (reducing and non-reducing Supplementary Fig. S3) and ESI-TOF mass spectrometry (Functional Genomics Center Zurich, University of Zurich, CH). Protein concentrations were determined by UV absorbance at 280 nm on a NanoDrop spectrophotometer ND-1000. Theoretical molar extinction coefficients were predicted using the ProtParam tool at the EXPASy proteomics server (<http://www.expasy.ch/tools/protparam.html>), and are listed in Supplementary Table SII.

#### Far-UV CD spectroscopy

CD measurements were performed on a J-710 spectropolarimeter (JASCO, USA) equipped with a Peltier temperature-controlling unit (PTC-348 WI), using a 0.1 cm path-length quartz cuvette containing 300 µl of sample. Proteins were concentrated to 12 µM in 20 mM potassium phosphate, 40 mM NaCl, pH 7.0. Far-UV spectra were recorded at 20 °C, in the range of 190–260 nm, with a scanning speed of 20 nm per min, in 0.5 nm increments, with 2 s response time and 2 nm bandwidth, as an averaged signal of 10 scans and finally corrected by subtraction of the background scan with buffer. Thermal unfolding transitions were measured by monitoring the CD signal at 210 nm as a function of increasing temperature, from 20 to 95 °C, at a heating rate of 1 °C/min, in 0.2 °C increments with 8 s response time and 2 nm bandwidth. The reproducibility of the thermal unfolding transitions was assessed by measuring in triplicates, resulting in apparent melting temperature (*T<sub>m</sub>*) values not differing by more than 0.2 K. For details concerning data analysis, see Supplementary Methods.

#### Carbohydrates

Ganglioside GD2 analog [Neu5Acα2-8Neu5Acα2-3(GalNAcβ1-4Galβ1-4Glc) was purchased from ELICITYL SA (Crolles, FR). 6'-sulfo sialyl Lewis<sup>X</sup> [Neu5Acα2-3(6-O-SO<sub>3</sub>)Galβ1-4(Fucα1-3)GlcNAcβ-O-(3-amino)propyl] glycoside was chemically synthesized by a synthesis that will be published elsewhere. Methyl sialyl Lewis<sup>X</sup> was purchased from Carbosynth Ltd. (Berkshire, UK). The identity and purity of the carbohydrates was checked using 2D NMR spectroscopy.

#### NMR spectroscopy

NMR experiments were conducted on Bruker AVANCE-III 500, 750 and 900 MHz NMR spectrometers (all equipped with a cryogenic probe head except for AVIII-750 MHz). 2D <sup>1</sup>H–<sup>15</sup>N HSQC<sup>2</sup> spectra were recorded at 20 °C with the following acquisition parameters: For Siglec-7: 2048 (t<sub>2</sub>) × 180 (t<sub>1</sub>) complex points in each dimension, with 4 scans, a spectral width of 13.9 ppm (<sup>1</sup>H) × 31.9 ppm (<sup>15</sup>N), at a carrier frequency centered at 118.4 ppm, at 900 MHz; for Siglec-8: 2048 (t<sub>2</sub>) × 256 (t<sub>1</sub>) complex points in each dimension, with 2 scans, a spectral width of 16.0 ppm (<sup>1</sup>H) × 48.0 ppm (<sup>15</sup>N), at a carrier

<sup>2</sup> Abbreviations used: 2D <sup>1</sup>H–<sup>15</sup>N HSQC, two-dimensional <sup>1</sup>H–<sup>15</sup>N heteronuclear single quantum coherence; Ig-like, immunoglobulin-like; sLe<sup>X</sup>, sialyl Lewis<sup>X</sup>; 6'S sLe<sup>X</sup>, 6'-sulfo sialyl Lewis<sup>X</sup>.

frequency centered at 119.0 ppm, at 500 MHz; for Siglec-9: 2048 (t<sub>2</sub>) × 180 (t<sub>1</sub>) complex points in each dimension, with 4 scans, a spectral width of 16.0 ppm (<sup>1</sup>H) × 31.9 ppm (<sup>15</sup>N), at a carrier frequency centered at 118.6 ppm, at 750 MHz. Protein samples were concentrated at 310 μM (Siglec-7), 300 μM (Siglec-8) and 360 μM (Siglec-9), in 20 mM potassium phosphate, 40 mM NaCl, pH 7.0 and 5% D<sub>2</sub>O. NMR titration experiments were performed by recording a series of 2D <sup>1</sup>H–<sup>15</sup>N HSQC spectra thereby monitoring changes in chemical shifts after stepwise addition of glycan ligand from stock solutions of ganglioside GD2 analog at 9.2 mM, 6'-sulfo sialyl Lewis<sup>x</sup> at 7.65 mM or methyl sialyl Lewis<sup>x</sup> at 9.6 mM. <sup>1</sup>H chemical shifts are referenced to 2,2-dimethyl-2-silapentane-5-sulfonic acid (DSS). <sup>15</sup>N chemical shifts are indirectly referenced using a scaling factor of 0.101329118 [17]. NMR data were processed in TopSpin 2.1 or 3.0 (Bruker BioSpin, CH) and further analyzed in Sparky (T.D. Goddard and D.G. Kneller, SPARKY 3, University of California, San Francisco, USA).

## Results

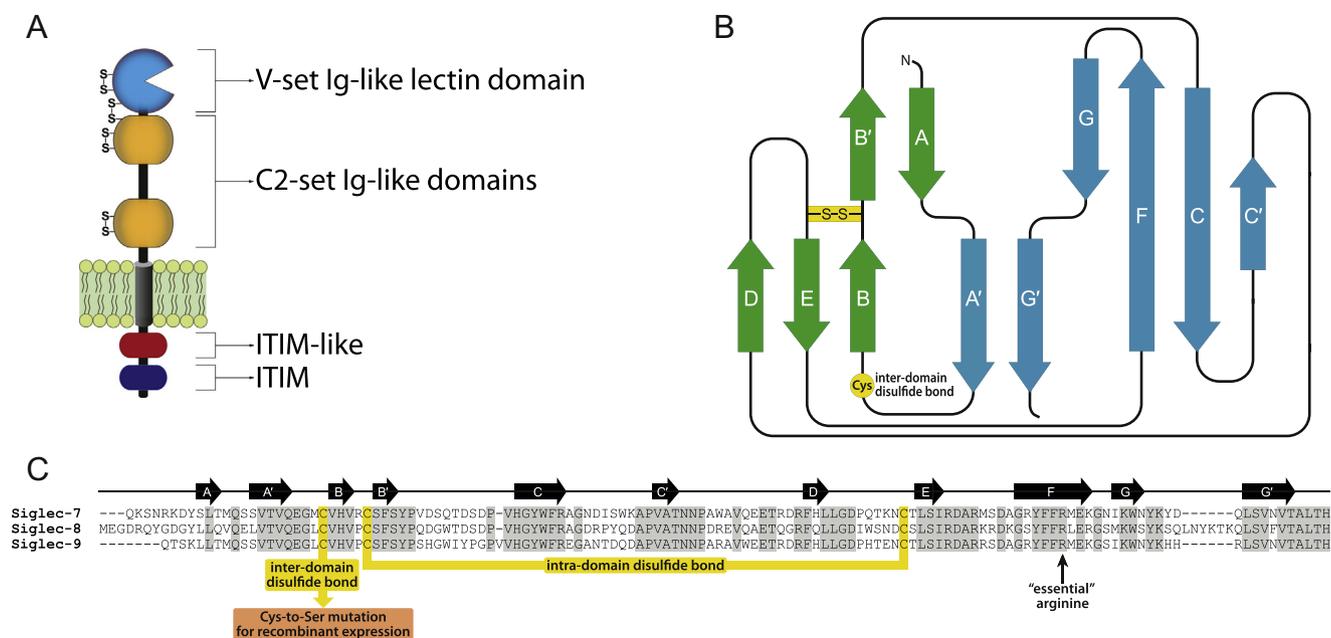
### Construction of expression plasmids

For the development of an efficient *E. coli*-based expression system for Siglec lectin domains, we selected the lectin domains of human Siglec-7, -8 and -9 as test subjects. These three Siglecs share a similar domain architecture, with an N-terminal sialic-acid binding V-set immunoglobulin-like domain, two C2-set immunoglobulin-like domains, a single transmembrane helical segment and immunoreceptor tyrosine-based inhibitory motifs (ITIM/ITIM-like) in their cytoplasmic tails (Fig. 1A). The typical folding topology of a Siglec lectin domain is displayed in Fig. 1B, that was derived from crystal structures of Siglec-5 and -7 (PDB entries 2ZG1 and 1O7V, respectively). Each lectin domain contains three highly conserved cysteine residues: two form an intra-domain disulfide bond and the third one is involved in an inter-domain disulfide bond with the subsequent C2-set domain (Fig. 1). For our expression constructs of the isolated lectin domains, the latter cysteine (Cys41, Cys42 and Cys36 in Siglec-7, -8 and -9, respectively) was mutated

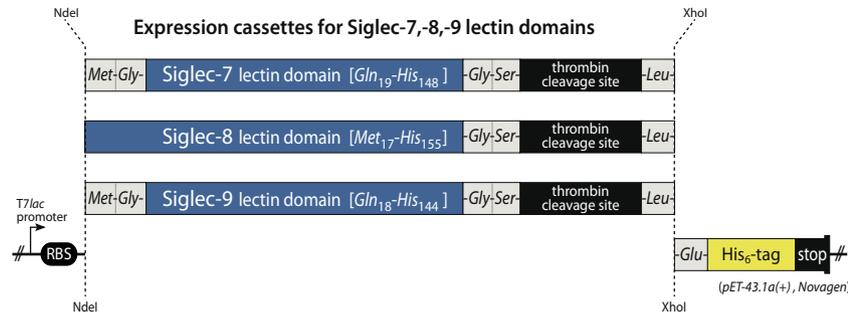
to a serine to eliminate a potential cause of misfolding due to wrong disulfide pairing (Fig. 1C). The sequences encoding for each Siglec lectin domain carrying the Cys-to-Ser point-mutation, followed by a two-residue spacer (GlySer) and a thrombin cleavage site (LeuValProArg↓GlySer) were cloned between NdeI and XhoI restriction sites of a pET-43.1(a) vector (Novagen). The resulting constructs, illustrated in Fig. 2, encode for a fusion protein consisting of the respective Siglec lectin domain, followed by a cleavable C-terminal hexahistidine-affinity-tag (His<sub>6</sub>-tag). The C-terminal position of the His<sub>6</sub>-tag was chosen to ensure that only fully translated proteins will be tagged and potentially occurring abortive products resulting from premature termination of translation, for example due to the high abundance of rare codons in the human cDNAs (Supplementary Table SIII), will not be purified by affinity chromatography. Nucleotide and amino acid sequences of the Siglec expression constructs are listed in the Supplementary Table SIV.

### Expression and purification of soluble Siglec lectin domains

To identify suitable expression conditions that produce soluble Siglec lectin domains, the Siglec-9 construct was subjected to analytical expressions in different *E. coli* host strains. The expression was first tested in BL21-CodonPlus(DE3)-RIL, a conventional strain that supplies additional tRNAs for four rare codons. The naturally reducing environment in the cytoplasm of this strain strongly disfavors the formation of disulfide bonds. Using these cells, Siglec-9 was only insolubly expressed at all tested culturing conditions (data not shown). We further tested the Siglec-9 expression in Rosetta-gami B(DE3), an engineered strain that provides a more oxidative cytoplasmic environment through inactivating mutations in its thioredoxin reductase (*trxB*) and glutathione reductase (*gor*) genes, thus allowing the formation of stable disulfide bonds. In addition, this strain produces extra tRNAs for six rare codons. Initial tests in Rosetta-gami B(DE3) cells revealed that the strain was capable of expressing soluble Siglec-9, however only with poor yields. To further improve the soluble expression of Siglec-9, we systematically investigated the effects of various expression parameters, including different induction temperatures, inducer



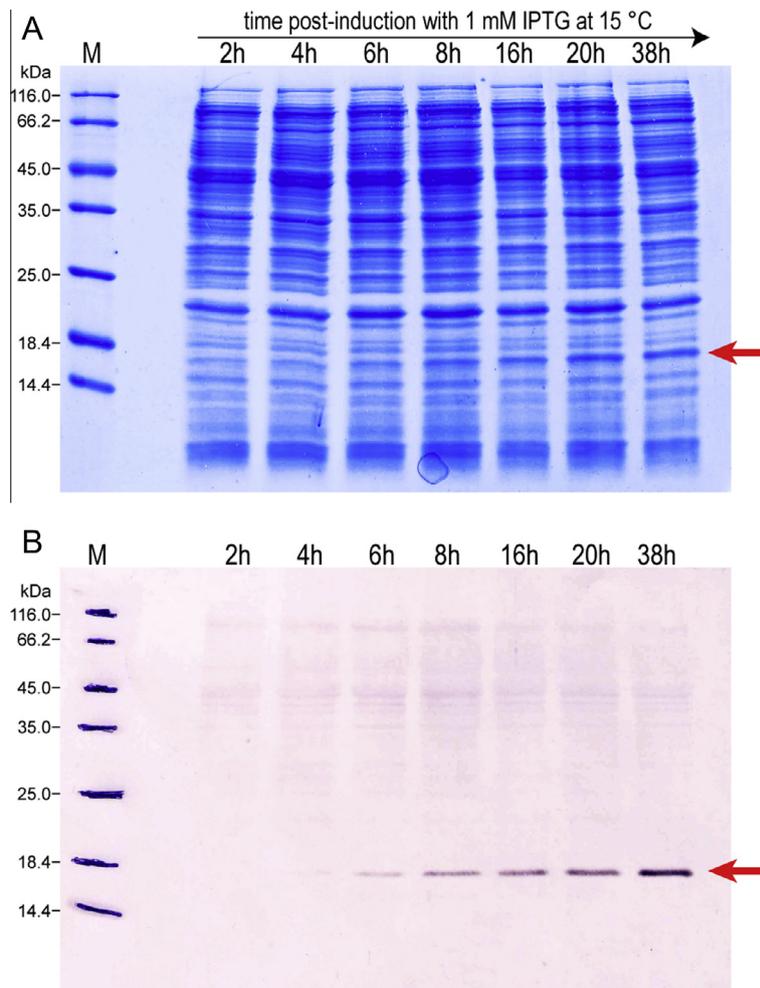
**Fig. 1.** (A) Scheme illustrating the molecular architecture of Siglec-7, -8 and -9 receptors. (B) Fold topology of a Siglec V-type Ig-like lectin domain (derived from Siglec-5 and -7 crystal structures, PDB entries 2ZG1 and 1O7V, respectively). (C) Sequence alignment of the mature protein sequences of human Siglec-7, -8, and -9 lectin domains.



**Fig. 2.** Expression constructs for the isolated lectin domains of human Siglec-7, -8 and -9.

concentrations and induction durations (Supplementary Fig. S1). As judged from SDS-PAGE and anti-His-tag Western Blot analysis of the soluble cell lysate fractions, the largest amount of soluble Siglec-9 was expressed at the lowest tested induction temperature of 15 °C, using an IPTG concentration of 1 mM and the longest tested post-induction period of 38 h (Fig. 3). These optimized expression conditions were applied for the preparative production of all three Siglec lectin domains. Cells cultured in LB medium or M9 minimal medium were harvested 40 or 60 h after induction, respectively.

The soluble proteins were purified using immobilized metal-ion affinity chromatography (Ni-NTA), followed by thrombin cleavage for His-tag removal. A final size exclusion chromatographic step yielded highly pure proteins, eluting as single, symmetrical peaks with retention volumes in the expected range for their calculated molecular weights, indicating a homogeneous conformation and monomeric state of the recombinant proteins (Fig. 4 and Supplementary Fig. S2). We reproducibly obtained yields of 1.3–1.5, 1.1–1.5 and 0.4–1.0 mg per liter of LB culture for purified, native Siglec-7, -8 and -9, respectively (each determined from a 1 L

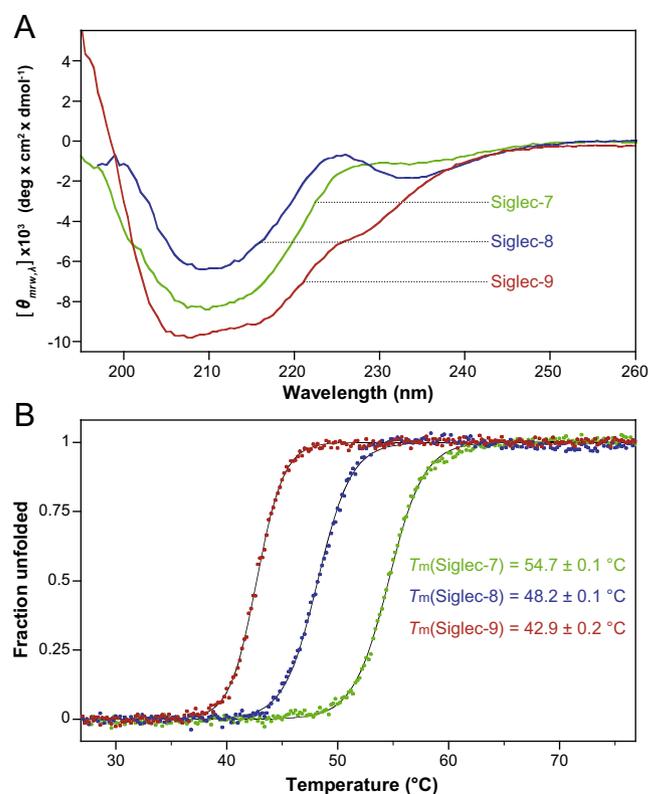


**Fig. 3.** Expression profiles of soluble Siglec-9 in Rosetta-gami B(DE3) at 15 °C induced with 1 mM IPTG. (A) Coomassie blue-stained reducing SDS-PAGE (15%) gel of soluble cell extracts obtained from different time-points after induction. (B) Western Blot analysis of a duplicate gel transferred onto nitrocellulose and probed by an anti-His-tag antibody for detection of Siglec-9. For better comparability, equal amounts of samples corresponding to an OD<sub>600</sub> of 0.2 of original cells were loaded. M, molecular weight markers; red arrows indicate the position of the Siglec-9 protein bands. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

expression volume). Similar yields were achieved from expressions in M9 minimal medium. Each protein appeared as a single band in SDS-PAGE under both reducing and non-reducing conditions (Supplementary Fig. S3), with the non-reduced protein bands migrating slightly faster, demonstrating the uniform presence of the intra-domain disulfide bond. The identity, purity and the presence of the intra-domain disulfide bond in the three recombinant proteins was additionally confirmed by electrospray ionization mass spectrometry (ESI-TOF-MS). The mass spectrometric analysis further revealed that the produced Siglec-7 and -9 lectin domains were devoid of their N-terminal initiator methionine residue, presumably through the activity of the cotranslationally acting *E. coli* methionine aminopeptidase [18], leaving a single glycine as the only non-natural N-terminal extension in both recombinant proteins. The observed molecular masses for Siglec-8 are in agreement with the expected ones, starting from its predicted natural N-terminal residue (Met17) of the mature protein sequence. (See Supplementary Fig. S4 for mass spectra and Supplementary Table SV for theoretical and measured molecular weights.)

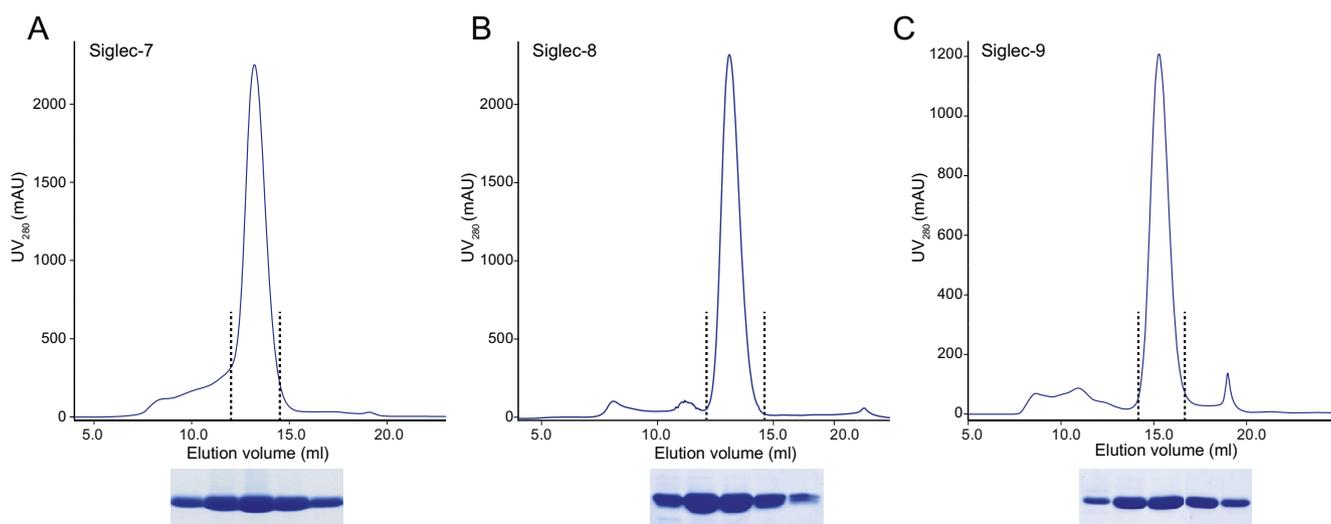
#### Characterization and functional activity of the recombinant Siglec lectin domains

The structural integrity of the three recombinant proteins was assessed by far-ultraviolet circular dichroism (CD) spectroscopy. The recorded far-UV CD spectra (Fig. 5A) feature typical traces of a well-folded protein whose secondary structure is dominated by  $\beta$ -sheets [19]. Each spectrum displays a prominent minimum at 210, 209.5 or 208 nm wavelength for Siglec-7, -8 and -9, respectively. The observed spectrum of Siglec-7 is consistent with previously published data [11]. To evaluate thermal stabilities of the recombinant lectin domains, temperature-induced unfolding transitions were recorded by following changes in the CD signal at 210 nm as a function of temperature. Each of the three proteins (Fig. 5B) was found to unfold in a highly cooperative, two-state ( $F \rightleftharpoons U$ ) transition, without indication for consecutive unfolding steps or the formation of any stable intermediates, demonstrating that the recombinant proteins are homogeneously folded with intact and stable tertiary structures. The apparent melting temperatures ( $T_m$ ) for Siglec-7, -8 and -9 were 55, 48 and 43 °C, respectively. Under the experimental conditions used, the unfolding transitions for all three proteins proved to be irreversible and were



**Fig. 5.** Far-UV CD spectroscopic analysis of the purified, recombinant Siglec-7, -8 and -9 lectin domains. All CD experiments were performed at protein concentrations of 12  $\mu$ M, in 20 mM potassium phosphate, 40 mM NaCl, pH 7.0. (A) Overlay of the far-UV CD spectra of Siglec-7 (green), Siglec-8 (blue) and Siglec-9 (red) recorded at 20 °C. (B) Thermal unfolding transitions of the three recombinant Siglec lectin domains. Apparent melting temperatures ( $T_m$ ) were extracted from fitting the data to a reversible two-state transition reaction (solid lines). The shown  $T_m$ -values and indicated errors are mean values from triplicate measurements and their respective standard deviations. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

accompanied by extensive insoluble aggregation. From the shown far-UV CD experiments we concluded that all three recombinantly expressed Siglec lectin domains were properly folded. Siglec-9 was the least temperature stable and we therefore carried out all bio-



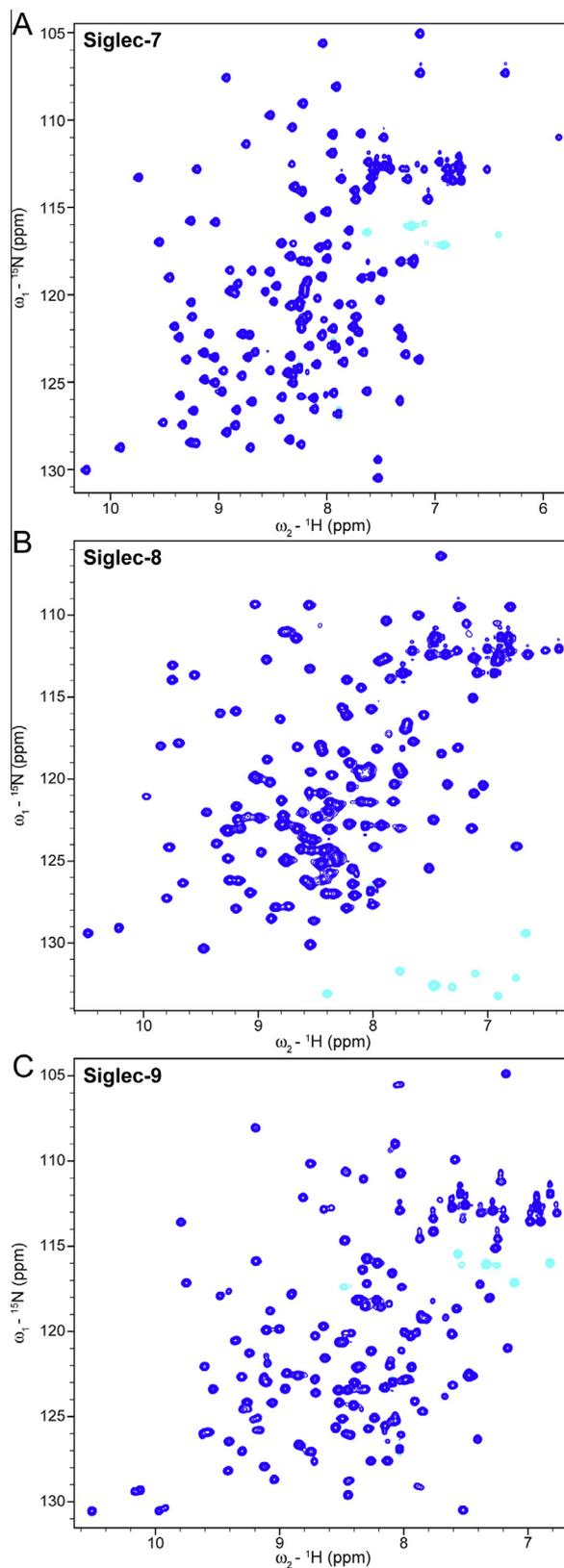
**Fig. 4.** Size exclusion chromatography elution profiles from purifications of recombinant (A) Siglec-7, (B) Siglec-8 and (C) Siglec-9 lectin domains, detected by UV absorbance at 280 nm wavelength throughout the run (solid lines). Samples from fractions corresponding to the major peaks (indicated by dotted lines in the chromatograms) are shown in sections of Coomassie blue-stained SDS-PAGE gels below.

physical measurements at 20 °C, well below the onsets of unfolding to prevent any irreversible aggregation.

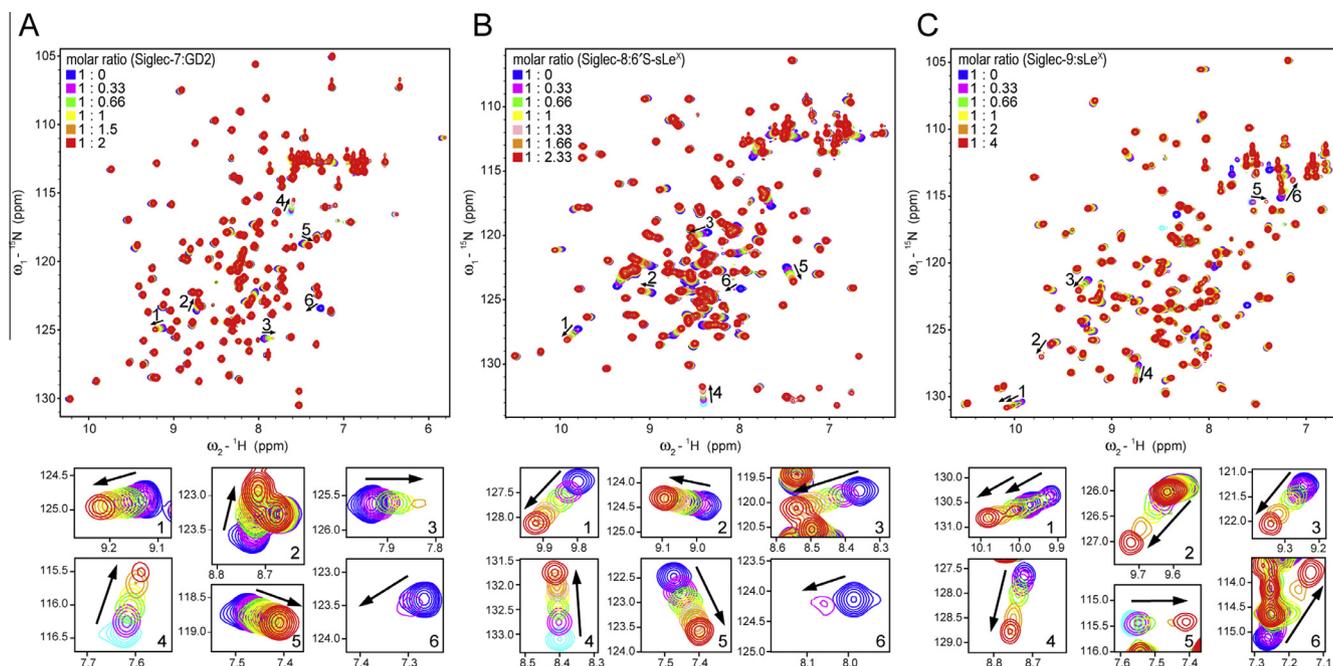
The presence of tertiary fold was further verified by two-dimensional NMR spectroscopy. The 2D  $^1\text{H}$ - $^{15}\text{N}$  heteronuclear single-quantum coherence (HSQC) NMR spectra of the uniformly  $^{15}\text{N}$ -labeled proteins showed excellent chemical shift dispersion, typical for well-folded  $\beta$ -sheet-rich proteins, with equal peak intensities and line widths expected for monomeric species (Fig. 6). For all three proteins the observed number of backbone amide resonances was consistent with those expected. To prove whether the recombinant proteins were also functionally active, NMR binding experiments with their proposed natural carbohydrate targets were performed. Titrations of ganglioside GD2 into Siglec-7, 6'-sulfo sialyl Lewis<sup>X</sup> into Siglec-8 and sialyl Lewis<sup>X</sup> into Siglec-9 resulted in numerous significant chemical shift changes in the 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra (Fig. 7). Upon stepwise addition of ligand, most of the affected peaks shifted gradually with constant signal intensities, suggesting that binding is characterized by fast chemical exchange in all three protein-carbohydrate complexes. Only few peaks appeared to follow a fast-to-intermediate exchange regime, indicated by partial or complete signal disappearance (e.g. signals 3 and 6 in Fig. 7 A, signal 6 in panel B and signals 5 and 6 in panel C). The NMR titration experiments clearly demonstrate that all three recombinant Siglec lectin domains are capable of binding their proposed carbohydrate ligands and are thus functionally active.

## Discussion

The lack of efficient and applicable systems for the recombinant production of homogeneous and active Siglec lectin domains has been a long-standing impediment for their detailed functional and structural investigation. All previous attempts to express these domains in *E. coli* have been hampered by the insolubility of the expressed proteins [11–13]. The presence of disulfide bonds in their native structures that cannot be stably formed in the reducing environment of the bacterial cytoplasm, certainly contributes to their high tendency to misfold and aggregate during expression. Complicated and laborious oxidative refolding procedures had to be applied to recover active protein from inclusion bodies. Expression in the naturally oxidizing *E. coli* periplasm is an often used alternative route to produce disulfide-containing proteins in a soluble and active state; however this method also suffers from severe drawbacks arising from inefficient secretion, proteolytic degradation or improper folding of the target proteins [20,21]. In recent years, a variety of strains has been specifically engineered for the cytoplasmic expression of disulfide-bonded proteins, based on the disruption of the thioredoxin reductase (encoded by the *trx*B gene) and/or the glutathione reductase (encoded by the *gor* gene) reductive pathways present in the bacterial cytoplasm [22–25]. Such strains have already been successfully implemented for the soluble expression of several other disulfide-containing heterologous proteins, for which conventional expression strains failed [26]. Surprisingly, their application as an expression host for Siglec lectin domains had never been explored before. In the present study, the *trx*B *gor* double mutant strain Rosetta-gami B(DE3) and the conventional strain BL21-CodonPlus(DE3)-RIL were evaluated for their capability to express the Siglec-9 lectin domain in a soluble and folded state. Our analysis revealed that BL21-CodonPlus(DE3)-RIL was not capable of producing soluble Siglec-9 at any condition tested. In contrast, small amounts of soluble Siglec-9 were found to be expressed in the oxidative cytoplasm of Rosetta-gami B(DE3) cells. Based on an extensive screening of growth and induction parameters, we successfully identified optimized conditions at which the soluble expression of Siglec-9 was



**Fig. 6.** 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of the free Siglec lectin domains ( $^{15}\text{N}$ -labeled) recorded at 20 °C. (A) Siglec-7, (B) Siglec-8 and (C) Siglec-9. Protein concentrations were 310  $\mu\text{M}$  (Siglec-7), 300  $\mu\text{M}$  (Siglec-8) or 360  $\mu\text{M}$  (Siglec-9) dissolved in 20 mM potassium phosphate, 40 mM NaCl, pH 7.0 and 5%  $\text{D}_2\text{O}$ , at  $^1\text{H}$  NMR frequencies of 900, 500, or 750 MHz, respectively. Positive cross-peaks are shown in blue and negative cross-peaks in cyan, corresponding to folded signals. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



**Fig. 7.** NMR titration experiments of the recombinant Siglec lectin domains with their proposed carbohydrate ligands: (A) Siglec-7 with ganglioside GD2 analog, (B) Siglec-8 with 6'-sulfo sialyl Lewis<sup>X</sup> and (C) Siglec-9 with sialyl Lewis<sup>X</sup>. Depicted are 2D <sup>1</sup>H–<sup>15</sup>N HSQC spectra of the free proteins colored in blue (positive cross-peaks) and cyan (folded cross-peaks), successively overlaid by the spectra collected after gradual addition of the individual ligands. Legends shown in the top left corners of each panel indicate the color codes used to display spectra recorded at different titration points and the corresponding molar ratios (protein vs. carbohydrate). For each titration, six representative peaks (numbered 1–6) showing significant chemical shift changes upon ligand binding are shown in enlarged sections below. Arrows indicate the direction of cross-peak movement. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

significantly improved. Our results indicate that growth temperatures had the strongest impact on improving the expression in Rosetta-gami B(DE3) cells. The largest increase in soluble Siglec-9 expression was clearly gained from lowering the induction temperature to 15 °C. Expression at decreased temperatures is a common strategy to improve the solubility and stability of target proteins in *E. coli* [27–29]. These effects can most likely be ascribed to the slowed transcription and translation rates, causing a condition at which the cellular protein concentration is ideally balanced with the slow process of oxidative protein folding. Also, at such conditions, the bacterial folding machinery is under a decreased work load, liberating more free chaperones to aid correct protein folding. Furthermore, due to the strong temperature dependence of hydrophobic interactions [30,31], the aggregation of folding intermediates is generally decreased at low temperatures. Despite the often observed benefits of low-level induction [32,33], considerably larger amounts of soluble Siglec-9 were obtained by using 1 mM rather than 0.1 mM IPTG, indicating that high-level induction at low temperature is optimal. It is noteworthy to mention, that the Rosetta-gami B cells exhibited remarkably slow growth rates, with doubling times of  $\approx 2$ –2.5 h at 37 °C before induction, and  $\approx 40$  h at 15 °C during induction, all measured in M9 minimal medium. Similar observations were made for expressions performed in LB medium with slightly faster growth rates. Nevertheless, the cells were growing constantly and no cell lysis was observed at any time point. Therefore, long post-induction periods of 40 and 60 h for expression in LB and M9 minimal media, respectively, were required for maximal yields of soluble Siglec-9. The slow growth of *trxB gor* suppressor strains has previously been reported and is believed to be a consequence of the engineered oxidative cytoplasm that also affects endogenous cellular functions, rather than due to cellular toxicity of the heterologous expressed protein [23,24]. It can be speculated that the slow growth even enhances the proper protein folding and disulfide bond formation

for reasons similar to those attributed to low induction temperatures.

The application of these optimized expression conditions for the preparative over-expression of the Siglec-7, -8 and -9 lectin domains, followed by a standard purification procedure, yielded milligram amounts of soluble, highly pure, homogeneously folded and functionally active protein samples for all three tested candidates. To the best of our knowledge, this is the first report of a successful soluble expression of Siglec lectin domains in a prokaryotic host system. The presented method enables a fast, convenient and scalable production of these three Siglec lectin domains in an active state and in quantities sufficient for extensive *in vitro* experimentation, without the need for denaturing and refolding. Their easy access will facilitate future structural studies for example by X-ray crystallography and also functional investigations like ligand binding experiments using techniques like isothermal titration calorimetry (ITC) or surface plasmon resonance (SPR). Efficient and cost-effective isotope-labeling has been achieved, a prerequisite for multidimensional NMR spectroscopy. The observed NMR fingerprint spectra displayed excellent signal dispersion, qualifying these recombinant protein preparations for residue-specific epitope mapping studies and three-dimensional structure determination. Moreover, the presented system provides a suitable platform to generate engineered variants of Siglec lectin domains, interesting for a broad range of applications, for example the preparation of Siglec chimeras with altered binding specificities and affinities [34] or the identification of mutations that increase protein stability. In analogy to fluorescently-labeled antibodies as described by Markiv et al. [35], our system could be useful for producing GFP fusion proteins that could serve as immunofluorescent reagents for glycan ligand detection in immunoassays or immunohistochemistry. Likewise, Siglec lectin domains could be fused to a biotinylation target sequence to obtain streptavidin-affine molecules suitable for immobilization or oligomerization, leading to

increased affinities through avidity effects. Such an approach using tetramerized Siglec-7 ectodomains as a tool for flow-cytometric cell sorting of specific subsets of lymphocytes, has recently been reported [36]. Furthermore, the recombinant Siglec lectin domains can be used for immunization to trigger the production of Siglec-specific antibodies that can be employed as probes for targeting Siglecs or Siglec expressing cells *in vitro* or *in vivo* [3,37].

## Conclusion

A simple and efficient recombinant expression and purification protocol for Siglec lectin domains has been established allowing to produce quantitative amounts of highly pure, homogeneously folded, non-glycosylated and functionally active protein samples from a soluble expression in the *E. coli* cytoplasm. Our method was successfully applied for the over-expression and purification of the isolated lectin domains of human Siglec-7, -8, and -9. We therefore expect that a similar construct design and expression setup is also applicable to other members of the Siglec family, but also to structurally related proteins. The easy access to recombinant Siglec lectin domains will strongly facilitate future structural and functional investigations and is paving the way for a broad range of biomedical applications, including high-throughput ligand screening or the development of diagnostic tools for detection of specific glycan epitopes on host cells or pathogens.

## Conflict of interest statement

None declared.

## Acknowledgments

The authors thank Daniel Perez for kindly providing Rosetta-gami B (DE3) cells for initial expression tests. Also, we thank Christoph Giese, Daniel Roderer and Maximilian Sauer for valuable advices about the CD instrumentation usage and CD data analysis. We are further grateful to Anne Diehl, Georg Dorn, Kyle Fowler and Sébastien Campagne for helpful discussions and critical reading of the manuscript. This work was supported by the Swiss National Science Foundation (Sinergia Grant Number: CRSI3\_127333) to F.H.-T.A. and B.E.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.pep.2015.01.005>.

## References

- [1] P.R. Crocker, J.C. Paulson, A. Varki, Siglecs and their roles in the immune system, *Nat. Rev. Immunol.* 7 (2007) 255–266.
- [2] A. Varki, T. Angata, Siglecs – the major subfamily of I-type lectins, *Glycobiology* 16 (2006) 1R–27R.
- [3] M.S. Macauley, P.R. Crocker, J.C. Paulson, Siglec-mediated regulation of immune cell function in disease, *Nat. Rev. Immunol.* 14 (2014) 653–666.
- [4] J. Müller, L. Nitschke, The role of CD22 and Siglec-G in B-cell tolerance and autoimmune disease, *Nat. Rev. Rheumatol.* 10 (2014) 422–428.
- [5] C. Jandus, K.F. Boligan, O. Chijioke, H. Liu, M. Dahlhaus, T. Demoulin, C. Schneider, M. Wehrli, R.E. Hunger, G.M. Baerlocher, H.U. Simon, P. Romero, C. Munz, S. von Gunten, Interactions between Siglec-7/9 receptors and ligands influence NK cell-dependent tumor immunosurveillance, *J. Clin. Invest.* 124 (2014) 1810–1820.
- [6] C. Jandus, H.U. Simon, S. von Gunten, Targeting Siglecs – a novel pharmacological strategy for immuno- and glycotherapy, *Biochem. Pharmacol.* 82 (2011) 323–332.
- [7] M.K. O'Reilly, J.C. Paulson, Siglecs as targets for therapy in immune-cell-mediated disease, *Trends Pharmacol. Sci.* 30 (2009) 240–248.
- [8] A.P. May, R.C. Robinson, R.T. Aplin, P. Bradfield, P.R. Crocker, E.Y. Jones, Expression, crystallization, and preliminary X-ray analysis of a sialic acid-binding fragment of sialoadhesin in the presence and absence of ligand, *Protein Sci.* 6 (1997) 717–721.
- [9] M.S. Alphey, H. Attrill, P.R. Crocker, D.M. van Aalten, High resolution crystal structures of Siglec-7. Insights into ligand specificity in the Siglec family, *J. Biol. Chem.* 278 (2003) 3372–3377.
- [10] M. Vinson, P.A. van der Merwe, S. Kelm, A. May, E.Y. Jones, P.R. Crocker, Characterization of the sialic acid-binding site in sialoadhesin by site-directed mutagenesis, *J. Biol. Chem.* 271 (1996) 9267–9272.
- [11] N. Dimasi, L. Moretta, R. Biassoni, R.A. Mariuzza, Expression, crystallization and preliminary crystallographic analysis of the extracellular IgV-like domain of the human natural killer cell inhibitory receptor p75/AIRM1, *Acta Crystallogr. D Biol. Crystallogr.* 59 (2003) 1856–1858.
- [12] M.A. Zhuravleva, K. Trandem, P.D. Sun, Structural implications of Siglec-5-mediated sialoglycan recognition, *J. Mol. Biol.* 375 (2008) 437–447.
- [13] A.W. Barb, X. Wang, J.H. Prestegard, Refolded recombinant Siglec5 for NMR investigation of complex carbohydrate binding, *Protein Expr. Purif.* 88 (2013) 183–189.
- [14] J. Sambrook, *Molecular cloning: a laboratory manual*, in: D.W. Russell (Ed.), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 2001.
- [15] M. Cai, Y. Huang, K. Sakaguchi, G.M. Clore, A.M. Gronenborn, R. Craigie, An efficient and cost-effective isotope labeling protocol for proteins expressed in *Escherichia coli*, *J. Biomol. NMR* 11 (1998) 97–102.
- [16] H. Towbin, T. Staehelin, J. Gordon, Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications, *Proc. Natl. Acad. Sci. U.S.A.* 76 (1979) 4350–4354.
- [17] J. Markley, A. Bax, Y. Arata, C.W. Hilbers, R. Kaptein, B. Sykes, P. Wright, K. Wüthrich, Recommendations for the presentation of NMR structures of proteins and nucleic acids – IUPAC-IUBMB-IUPAB Inter-Union Task Group on the Standardization of Data Bases of Protein and Nucleic Acid Structures Determined by NMR Spectroscopy, *J. Biomol. NMR* 12 (1998) 1–23.
- [18] F. Frottin, A. Martinez, P. Peynot, S. Mitra, R.C. Holz, C. Giglione, T. Meinel, The proteomics of N-terminal methionine cleavage, *Mol. Cell. Proteomics* 5 (2006) 2336–2349.
- [19] S.M. Kelly, T.J. Jess, N.C. Price, How to study proteins by circular dichroism, *Biochim. Biophys. Acta* 1751 (2005) 119–139.
- [20] N. Ke, M. Berkman, Production of disulfide-bonded proteins in *Escherichia coli*, *Curr. Protoc. Mol. Biol.* 108 (2014) 16 11B 11–16 11B 21.
- [21] S. Schlegel, E. Rujas, A.J. Ytterberg, R.A. Zubarev, J. Luirink, J.W. de Gier, Optimizing heterologous protein production in the periplasm of *E. coli* by regulating gene expression levels, *Microb. Cell Fact.* 12 (2013) 24.
- [22] A.I. Derman, W.A. Prinz, D. Belin, J. Beckwith, Mutations that allow disulfide bond formation in the cytoplasm of *Escherichia coli*, *Science* 262 (1993) 1744–1747.
- [23] W.A. Prinz, F. Aslund, A. Holmgren, J. Beckwith, The role of the thioredoxin and glutaredoxin pathways in reducing protein disulfide bonds in the *Escherichia coli* cytoplasm, *J. Biol. Chem.* 272 (1997) 15661–15667.
- [24] P.H. Bessette, F. Aslund, J. Beckwith, G. Georgiou, Efficient folding of proteins with multiple disulfide bonds in the *Escherichia coli* cytoplasm, *Proc. Natl. Acad. Sci. U.S.A.* 96 (1999) 13703–13708.
- [25] J.C. Bardwell, K. McGovern, J. Beckwith, Identification of a protein required for disulfide bond formation *in vivo*, *Cell* 67 (1991) 581–589.
- [26] A. de Marco, Strategies for successful recombinant expression of disulfide bond-dependent proteins in *Escherichia coli*, *Microb. Cell Fact.* 8 (2009) 26.
- [27] C.H. Schein, M.H.M. Noteborn, Formation of soluble recombinant proteins in *Escherichia coli* is favored by lower growth temperature, *Nat. Biotech.* 6 (1988) 291–294.
- [28] H.P. Sorensen, K.K. Mortensen, Soluble expression of recombinant proteins in the cytoplasm of *Escherichia coli*, *Microb. Cell Fact.* 4 (2005) 1.
- [29] S. Sahdev, S.K. Khattar, K.S. Saini, Production of active eukaryotic proteins through bacterial expression systems: a review of the existing biotechnology strategies, *Mol. Cell. Biochem.* 307 (2008) 249–264.
- [30] R.L. Baldwin, Temperature dependence of the hydrophobic interaction in protein folding, *Proc. Natl. Acad. Sci. U.S.A.* 83 (1986) 8069–8072.
- [31] J.A. Schellman, Temperature, stability, and the hydrophobic interaction, *Biophys. J.* 73 (1997) 2960–2964.
- [32] M.J. Weickert, D.H. Doherty, E.A. Best, P.O. Olins, Optimization of heterologous protein production in *Escherichia coli*, *Curr. Opin. Biotechnol.* 7 (1996) 494–499.
- [33] F. Baneyx, M. Mujacic, Recombinant protein folding and misfolding in *Escherichia coli*, *Nat. Biotechnol.* 22 (2004) 1399–1408.
- [34] T. Yamaji, T. Teranishi, M.S. Alphey, P.R. Crocker, Y. Hashimoto, A small region of the natural killer cell receptor, Siglec-7, is responsible for its preferred binding to alpha 2,8-disialyl and branched alpha 2,6-sialyl residues. A comparison with Siglec-9, *J. Biol. Chem.* 277 (2002) 6324–6332.
- [35] A. Markiv, R. Beatson, J. Burchell, R.V. Durvasula, A.S. Kang, Expression of recombinant multi-coloured fluorescent antibodies in *gor*<sup>+</sup>/*trxB*<sup>-</sup> *E. coli* cytoplasm, *BMC Biotechnol.* 11 (2011) 117.
- [36] F. Gieseke, P. Mang, S. Viebahn, I. Sonntag, A. Kruchen, A. Erbacher, M. Pfeiffer, R. Handgretinger, I. Müller, Siglec-7 tetramers characterize B-cell subpopulations and leukemic blasts, *Eur. J. Immunol.* 42 (2012) 2176–2186.
- [37] C.L. Erickson-Miller, S.D. Freeman, C.B. Hopson, K.J. D'Alessio, E.I. Fischer, K.K. Kikly, J.A. Abrahamson, S.D. Holmes, A.G. King, Characterization of Siglec-5 (CD170) expression and functional activity of anti-Siglec-5 antibodies on human phagocytes, *Exp. Hematol.* 31 (2003) 382–388.