Structure–Activity Relationship of the Dimeric and Oligomeric Forms of a Cytotoxic Biotherapeutic Based on Diphtheria Toxin

Marcin Mielecki [†], Marcin Ziemniak [†], Magdalena Ozga, Radosław Borowski, Jarosław Antosik, Angelika Kaczyńska and Beata Pająk *

WPD Pharmaceuticals, Żwirki and Wigury 101, 02-089 Warsaw, Poland
* Correspondence: beata.pajak@wpdpharmaceuticals.com
† These authors contributed equally to this work.

Abstract: Protein aggregation is a well-recognized problem in industrial preparation, including biotherapeutics. These low-energy states constantly compete with a native-like conformation, which is more pronounced in the case of macromolecules of low stability in the solution. A better understanding of the structure and function of such aggregates is generally required for the more rational development of therapeutic proteins, including single-chain fusion cytotoxins to target specific receptors on cancer cells. Here, we identified and purified such particles as side products of the renaturation process of the single-chain fusion cytotoxin, composed of two diphtheria toxin (DT) domains and interleukin 13 (IL-13), and applied various experimental techniques to comprehensively understand their molecular architecture and function. Importantly, we distinguished soluble purified dimeric and fractionated oligomeric particles from aggregates. The oligomers are polydisperse and multimodal, with a distribution favoring lower and even stoichiometries, suggesting they are composed of dimeric building units. Importantly, all these oligomeric particles and the monomer are cystine-dependent as their innate disulfide bonds have structural and functional roles. Their reduction triggers aggregation. Presumably the dimer and lower oligomers represent the metastable state, retaining the native disulfide bond. Although significantly reduced in contrast to the monomer, they preserve some fraction of bioactivity, manifested by their IL-13RA2 receptor affinity and selective cytotoxic potency towards the U-251 glioblastoma cell line. These molecular assemblies probably preserve structural integrity and native-like fold, at least to some extent. As our study demonstrated, the dimeric and oligomeric cytotoxin may be an exciting model protein, introducing a new understanding of its monomeric counterpart's molecular characteristics.

Keywords: biotherapeutics; cytotoxin; diphtheria toxin; IL-13; inclusion bodies; refolding; disulfide bond; protein oligomerization; MALS; SAXS; LC/MS

1. Introduction

Immunotoxins (IT) are fusion biotherapeutics composed of two major parts: a receptor-binding moiety and the active toxic payload. The binding part is usually an antibody or a ligand directed towards a specific receptor expressed on the cell membrane, such as IL-13RA2. The active part of the IT is the toxic payload, such as drugs, radioisotopes, toxins, and enzymes [1]. Compared to most drugs, toxins act catalytically, do not develop drug resistance, and can be applied to both dividing and quiescent cells [2]. A major advantage of toxins, when compared to radionucleotides and many small molecule drugs, is a lack of unspecific toxicity to surrounding cells and easier handling. The latter is the main drawback of radiopharmaceuticals [3]. The most commonly used toxins are of plant origin, for example, gelonin and ricin [4], or from bacteria, such as diphtheria toxin (DT) and exotoxin A (PE), produced by *Corynebacterium diphtheriae* and *Pseudomonas aeruginosa*, respectively [5]. Both DT and PE have shown to be highly efficient in killing eukaryotic cells. These molecules can irreversibly modify the mammalian elongation factor

2 (eEF2), thus inhibiting protein translation and inducing cell death [1]. Numerous genetic engineering studies showed that the most potent and non-immunogenic are DT390 and PE38 variants [6]. Hence, DT was selected for the generation of the first immunotoxin, Denileukin Diftitox (Ontak), which was designed to target the IL-2 receptor, which is highly overexpressed in cutaneous T-cell lymphoma (CTCL) [7]. Since then, several toxin-based biotherapeutics directed toward lymphoma and leukemia-specific markers, including CD3, CD25, and CD38, have been introduced. They can be directly injected into the bloodstream, easily reaching the transformed target cells [8,9]. Solid tumors could also be treated with ITs, albeit with lower efficacy, caused mainly by the partial penetration within the tumor mass [10]. Examples of ITs successfully introduced in solid tumor therapy are TE-38, NBI-3001, and Cintredekin besudotox (CB) used in anti-glioblastoma multiforme (GBM) therapy. All of them are fusion proteins containing the PE catalytic domain directed to GBM-specific IL-4R (NBI-3001), transforming growth factor alpha (TGFα) (TE-38), or IL-13RA2 (CB) receptors [11]. In the case of GBM, the blood-brain barrier (BBB) is the major obstacle in systemic drug delivery. Thus, local intra-tumoral immunotoxin administration with the convection-enhanced delivery (CED) method is commonly used [12]. Based on previous results showing the high efficacy of CED-delivered mutated DT390-IL-13 [13], we are currently developing an efficient large-scale production method of the mutated DT390-IL-13 cytotoxin.

Biopharmaceuticals represent one of the most promising frontiers in therapy significant, yet some challenges concerning their production and handling need to be overcome. The living systems producing biotherapeutics are sensitive to minor changes in the production process; hence, even slight differences may significantly affect the biological activity of the therapeutics. Labile biomolecules, including immunotoxins, are prone to degradation when exposed to environmental factors, such as elevated temperature, extreme pH, freezing stresses, organic and inorganic compounds, shear force, and light exposure [14,15]. One of the most common and troublesome effects of external stress is protein aggregation, which can occur at all stages of protein product development, usually leading to reduced bioactivity or increased immunogenicity risk [16,17]. Most pharmaceutically viable proteins must be folded to maintain their activity and molecular crowding, typical for the highly concentrated industrial formulation, is one of the major factors responsible for protein misfolding and aggregation [18,19]. The intrinsic properties of some proteins may also trigger the formation of aggregates. However, factors related to the expression and the purification conditions are believed to play a vital role in this process [20,21]. Reversible aggregation is a significant benefit to product quality in some cases. The native oligomers have higher stability and solubility than their monomeric counterparts, and insulin may be regarded as a canonical example [22]. Nonetheless, in most cases, aggregation, and to some extent oligomerization, is considered a potential detriment to product quality and must be avoided during the manufacture and storage. Although oligomers are common side-products in industrial biotechnology, there are few reports investigating pharmaceutically relevant proteins, and most studies are dedicated to antibodies. For instance, it was reported that monoclonal antibodies (mAb) dimers might form compact or elongated structures, either covalent or non-covalent [23]. Noticeably, some dimeric monoclonal antibodies were reported to be no more immunogenic than unstressed monomers [24].

The oligomerization process of cysteine (Cys)-containing proteins may be promoted or stabilized by forming intermolecular disulfide linkages. In most cases, disulfide bonds stabilize protein structure, and their rearrangement or mis-assembly may destabilize a protein's structural integrity, leading to the loss of its physiological activity. The dissociation energy of disulfide bond is relatively low and may be further reduced by local peptide flexibility, solvent exposure, and chemical environment facilitating their rearrangements [25]. Despite its reversibility, the formation of non-native cystine linkages often causes abundant oligomerization and aggregation, which has a detrimental effect on the structure and function [26]. Moreover, disulfide shuffling with non-native bonds and free cysteines was also reported for the isolated, stored, or heat-treated bovine serum albumin

(BSA) resulting in the formation of dimers and oligomers [27]. It was found that cystine crosslinks largely contribute to forming BSA aggregates, and sulfhydryl interchange reactions probably contributed more to polymerization than sulfur oxidation. Thus, disulfide network rearrangement, also for "structural" linkages, may be more common than expected. Interestingly, the success of attempts to stabilize protein expression and refolding by mutating disulfide-involved cysteines depends on the particular residue and may have both positive and negative effects [28].

Herein we describe the characteristic of the mutated DT390-IL-13 cytotoxin aggregation identified during *E. coli*-based recombinant protein production process and its importance for its biological activity. The exploration of protein-protein interactions background is a fundamental approach to controlling the quality of cytotoxin production and further success of its development as a biopharmaceutic. Using size-exclusion chromatography (SEC) coupled to solution X-ray scattering (SAXS) (SEC-SAXS), we also characterize a similar construct composed of exotoxin A and IL-13 (for brevity, called Exo in the rest of the article) to compare its behavior in the solution as a reference point, since this fusion protein does not incline to oligomerize during purification.

2. Materials and Methods

2.1. Reagents

Luria-Bertani broth (LB) and MacConkey broth (MB) media, Triton X-100, 5 M NaCl solution, guanidinium hydrochloride, urea, 2,2′,2″,2‴-(Ethane-1,2-diyldinitrilo)tetraacetic acid (EDTA) (0.5 M, pH 8.0), D-(+)-Glucose, protein marker V and peqGOLD were purchased from VWR Chemicals (Gdańsk, Poland). Magnesium sulphate, Isopropyl β-D-1-thiogalactopyranoside (IPTG) and phosphate-buffered saline (PBS), 10× Solution (pH 7.4) were purchased from Alfa Aesar (Kandel, Germany). L(+)-Arginine, L(−)-Glutathione (oxidized), 1,4-Dithioerythritol (DTE) ≥ 99% were purchased from Acros Organics (New Jersey, NJ, USA). Egg white lysozyme, ampicillin sodium salt, Trizma hydrochloride solution (1 M), chromogenic horseradish peroxidase (HRP) substrate, 2,2′-azino-bis(3-ethylbenzothiazoline-6-sulfonic acid (ABTS), 30% hydrogen peroxide solution and citric acid monohydrate were purchased from Merck Life Science (Kenilworth, NJ, USA). Run-Blue TEO-Tricine Sodium Dodecyl Sulfate (SDS) Gel 4–12%, RunBlue 20×TEO-Tricine-SDS Running Buffer, RunBlue LDS Sample Buffer 4×, RunBlue 10× Sample Reducer, and InstantBlue Protein Stain were obtained from Expedeon Ltd. (Cambridge, UK). AktaPure 25 FPLC system and prepacked chromatography columns were purchased from Cytiva Life Sciences (Global Life Sciences Solutions, Warsaw, Poland). MTS (CellTiter 96® AQueous One Solution Cell Proliferation Assay) was purchased from Promega (Madison, WI, USA). Cell culture medium, fetal bovine serum (FBS) and penicillin/streptomycin antibiotic mixture were purchased from Biowest (Nuaillé, France). Gentamycin solution was from purchased from TOKU-E (Sint-Denijs-Westrem, Belgium) and albumin from human serum (HSA) was obtained from Merck Life Science (Kenilworth, NJ, USA).

2.2. Plasmid and Host Strain

The sequence coding for a single-chain protein consisting of two N-terminal diphtheria toxin domains (DT390) fused to IL-13 was expressed under T7 constitutive promoter in plasmid pWD-MCS in strain BL21(DE3) (New England Biolabs, Ipswich, MA, USA). The expression plasmid was constructed in the Waldemar Dębiński lab (Wake Forest University Health Sciences, Winston-Salem, NC, USA) (details can be found here [29]). Sequence identity of the full-length plasmid was confirmed by MiSeq Illumina sequencing (Genomed SA, Warsaw, Poland). The expression plasmid was propagated in DH5α strain (New England Biolabs, Ipswich, MA, USA). Purity and concentration of the protein preparations were determined spectrophotometrically using DeNovix DS-11 FX, (DeNovix, Wilmington, DE, USA). The recombinant protein consists of 506 amino acids, including initial methionine, with predicted M_w of 55156 Da, pI = 5.67, and $\varepsilon = 51715$ M^{-1} cm^{-1} at 280 nm, assuming all pairs of cysteine residues form cystines.

2.3. Transgene Expression and Purification of the Recombinant Cytotoxin from Inclusion Bodies

The cytotoxin fusion protein was produced in the T7 expression system with BL21(DE3). Clone selection was performed, and expression kinetics was analyzed to identify the best expresser and optimal post-induction time point (2 h). Bacterial biomass was produced in 5 L shaker flasks by fermentation cultures in the LB Miller Broth media supplemented with 100 mM sodium phosphate buffer pH 7.0, 0.4% glucose, 1 mM MgSO$_4$, and 0.2 mg/mL ampicillin. To increase the efficiency of protein production, fed-batch high cell density cultivation was implemented in 5 L Biostat A MO 5 L UniVessel Glass bioreactor (Sartorius Stedim, Kostrzyn, Poland) with initial glucose concentration increased to 1% and 20% glucose as a feeding solution at the induction point in the following conditions, T = 37 °C, pH above 7.0 (adjusted with 0.5 M NaOH), and dissolved oxygen above 25%. Expression was induced at the mid-exponential growth phase with 1 mM IPTG. Inclusion bodies were extracted by suspending the cell pellet in a buffer containing: 0.2 mg lysozyme, 1% Triton X-100, and 0.5 M NaCl. Washed inclusion bodies were solubilized in the buffer: 100 mM Tris-HCl pH 7.4, 7 M guanidinium hydrochloride, 2 mM EDTA. DTE was added to the final concentration of 65 mM. Protein refolding was initialized by a rapid dilution method with disulfide shuffling into the buffer: 100 mM Tris-HCl pH 7.4, 2 mM EDTA, 0.6 M arginine, 0.9 mM oxidized glutathione and continued by dialysis into 20 mM Tris-HCl pH 8.0 with 100 mM urea. The recombinant protein was purified by the two-step chromatography on AKTA Pure 25 M1 FPLC system with Unicorn 7.5.0.1460 software. Dialysate clarified by the high-speed centrifugation was separated by the anion-exchange chromatography in the HiTrap Q HP (5 mL) column with step gradient and 20 mM Tris-HCl pH 8.0—1 M NaCl buffer system. The protein was further purified by the size-exclusion chromatography in the Superdex 200 Increase 10/300 GL column, and stored at −80 °C.

2.4. Reducing and Non-Reducing SDS-Polyacrylamide Gel Electrophoresis (PAGE)

Samples denatured in the LDS Sample Buffer (4×) without heating for non-reducing (nrSDS) gels and with a reducing agent and heating for 5 min, at 95 °C, for reducing (rSDS) gels, were loaded into RunBlue precast 4–12% gradient gels. Gels were resolved in TEO-Tricine-SDS running buffer and stained with the InstantBlue protein stain. SDS-PAGE gel scans were analyzed in ChemiDoc MP Imaging Systems with ImageLab 5.2.1 software (Bio-Rad, Warsaw, Poland). The semi-quantitative densitometric analysis assessed the purity of the protein preparations. The apparent $M_\mathbf{w}$ of the selected bands was estimated by the point-to-point (semi-log) regression method. To assess stoichiometry of the resolved oligomeric bands, regression analysis of the measured relative front (R_f) parameter was also performed. The plot of the R_f normalized to R_f of the lowest standard band (36 kDa) against the logarithm of standard molecular masses (fitted to the power function of the form Ax^B where A and B are model parameters) was used to calculate estimated oligomeric stoichiometries. R_f-estimated stoichiometries were compared to the theoretical stoichiometries by the one-sample *t*-test in the GraphPad Prism 9.3.1 software (LLC, San Diego, CA, USA). A weighted linear least-squares regression of the correlation of the assumed and predicted stoichiometries were also calculated.

2.5. Molecular Dynamic Simulations

All protein models were based on two crystallographic structures deposited in the PDB, 1F0L for DT390 and 3LB6 for IL-13. A missing loop in the 1F0L structure between cysteines 187 and 202 was built in YASARA 13.6.16 software. To probe the energy landscape of the fusion protein, five different arbitrary promodels were generated by changing selected dihedral angles in the interdomain linker. The promodels were subjected to molecular dynamics simulation in the NPT ensemble in YASARA (accessed 2 July 2022). Initial cell dimensions were 100/120/100 Å with periodic boundaries. The cell was filled with water with the TIP3P explicit solvent model and neutralized with NaCl. Protonation states for Asp, Glu, His, Lys, and Cys were assigned at pH 7.4 using Ewald summation to compute electrostatic potential [30]. Yamber3 force field [31] with 7.86-A cutoff for

non-bonded interactions was applied at T = 298 K. Particle Mesh Ewald algorithm was used for the long-range electrostatic interactions, and pressure was controlled with the water density set to 0.997 g·mL^{-1} with rescaling simulation cell dimensions. DT enzymatic and translocation domains atoms were fixed. The structures were energy-minimized by the steepest descent minimization over 50 ps. The procedure continued by the rescale-atom-velocities temperature control with a Berendsen thermostat and 2 fs time step until at least 20 ns total time was reached. Interdomain energy between DT390 and IL-13 was computed with BindEnregyObj (accessed 2 July 2022) that operates by subtracting the energy of the bound state from the energy of separated parts at infinite distance (the unbound state). The more positive the binding energy, the more favorable the interaction in the context of the chosen force field. The resulting energy, a sum of the potential force field energy components, does not include the solvation effects and ignores entropic terms. Thus, it is only a qualitative estimate to compare the stability of these structures. The hydrodynamic and structural properties were computed in the HullRad 8 software (accessed 2 July 2022) [32], applying full-atom models.

2.6. Batch Dynamic Light Scattering (DLS)

Size distributions for the oligomeric fractions were determined by the batch-mode DLS experiments conducted on DynaPro NanoStar 18.5 using auto-attenuation mode (Wyatt Technology Corporation, Santa Barbara, CA, USA). At least 10 isothermal (T = 25 °C) measurements, each with 10 acquisition scans for 5 or 10 s were recorded and analyzed in Dynamics 7.1.9.3 (accessed 2 July 2022) software (Wyatt Technology Corporation, Santa Barbara, CA, USA) using following parameters: PBS refractive index = 1.333, viscosity = 1.019 cP (20 °C), dn/dc = 0.185 mL·g^{-1}. Protein samples in PBS buffer pH 7.4 in the concentration range from 0.8 to 3.3 mg·mL^{-1} were centrifuged before collecting DLS data and were measured in plastic cuvettes with 2 mm optical path length. Dispersity and measurement quality was verified with bovine serum albumin solution in 0.9% NaCl as a monomer control and reference size. Translational diffusion coefficient (D_t), average translational hydrodynamic radius (R_h), and polydispersity parameters were determined using a built-in cumulants algorithm. The polydispersity divided by the estimated R_h was expressed as the polydispersity percent parameter. A hard-sphere model for globular proteins was applied to estimate the weight-averaged molar mass. Cumulative mass- or intensity-weighted size distributions with resolved particle populations were displayed on the regularization graphs with a non-negative least squares (NNLS) fitting algorithm (DynaLS).

2.7. Size-Exclusion Chromatography Coupled to Multi-Angle Light Scattering and DLS (SEC-MALS-DLS)

The protein samples were filtered and separated at 4 °C on Superose 6 Increase 10/300 GL column in PBS buffer pH 7.4 using NGC Chromatography System (Bio-Rad, Warsaw, Poland) coupled to scattering and quasi-elastic light scattering module (QELS, θ = 90°) detectors MiniDAWN Treos (Wyatt Technology Corporation, Santa Barbara, CA, USA) and refractometer ERC RefractoMax520 (DataApex, Prague, The Czech Republic). QELS temperature probe was enabled using the following parameters: PBS refractive index = 1.331, viscosity = 0.890 cP (25 °C), and dn/dc = 0.185 mL·g^{-1}. All data were collected and analyzed in ChromLab 6.0.0.34 (Bio-Rad, Warsaw, Poland) and ASTRA 7.1.2.5 (Wyatt Technology Corporation, Santa Barbara, CA, USA) software (accessed 2 July 2022). BSA solution in 0.9% NaCl served as a molecular standard for detector normalization, peak alignment and band broadening correction. Debye model with the $R_\Phi/K^*c = MP_\Phi$ formalism and fitting the plot of $\sin^2(\Phi/2)$ against R_Φ/K^*c was applied for the data processing of the chromatogram sections. The QELS autocorrelation data were fitted to the single exponential decay. Both R_h and M_w were derived for monomeric and dimeric peaks approximately at their full width at half maximum (FWHM) to minimize the variance of non-homogenous fractions. Average M_w, M_w distributions, polydispersity parameters (M_w/M_n, M_z/M_n), D_t, R_h, and

the cumulative and differential weight fractions distributions of the molar mass were also determined.

2.8. Size-Exclusion Chromatography Coupled to Small-Angle X-ray Scattering (SEC-SAXS)

All samples were passed continuously through the in-line SEC system (Agilent 1200 HPLC) connected to Superose 6 Increase 3.2/600 column (GE Healthcare, Chicago, IL, USA) and PBS buffer (pH = 7.4) using a flow rate 0.075 mL/min. X-ray scattering measurements were conducted at the B21 beamline of Diamond Light Source (Didcot, UK) using 12.4 keV X-rays at 75×75 μm^2 beam-size with 0.8×2 mm photon cross-section and flux 3.0 1012 photon/s, and approximately 620 frames (exposure time = 3 s, *q* range from 0.0031 to 0.38 Å^{-1}) were collected per sample using Eiger 4M detector. Initial SEC-SAXS data processing was performed using BioXTAS RAW 2.11 program [33,34]. At least 100 buffer profiles were averaged for background subtraction, and sample profiles were merged for each sample. Single Value Decomposition coupled with Evolving Factor Analysis (SVD-EFA) analyses were performed for each sample after subtraction. All scattering profiles were uncertainty-weighted and stored as column vectors in the data matrix A. The range of each potentially overlapping peak in the data was determined using EFA [35]. Rapid increases in the rank of submatrices A$_i$ are linked to the elution of new components; thus, the start of each peak was identified by plotting of SVs of the initial *n* columns as a function of *n* (which corresponds to the time evolution of an SV during the elution, forward EFA plot). Similarly, a sudden decrease in the ranks of an A$_i$ indicates the end of a given peak, which can be found by plotting SVs of the *n* columns removed from A, also a function of *n* (time-reversal evolution of SVs, backward EFA plot). Pairs of points indicating abrupt changes in the ranks of A$_i$ from forwards and backward EFA plots allow determining concentration windows for a given component. To calculate final *I*(*q*) profiles for each SV, the matrix of basis vectors U is non-orthogonally rotated by matrix R, giving a concentration of each component (c$_i$ column of concentration matrix C), c$_i$ = Ur$_i$.

2.9. SAXS Structural Modeling

R$_g$, *M*$_w$, and *D*$_{max}$ values and *P*(*r*) functions were calculated using BioXTAS RAW directly or via the ATSAS package (version 3.0.5) using RAW as a GUI (accessed 2 July 2022) [34,36]. *P*(*r*) functions were calculated automatically using the BIFT algorithm without modifications [37]. In the Guinier approximation, *R*$_g$ was calculated automatically and then tailored to reduce systematic deviation in the distribution of normalized residues, and *R*$_g$ based on *P*(*r*) was extracted automatically after computation. Methods of molecular weight calculations: (i) from Porod volume via SAXSMoW 2.0 (accessed 2 July 2022) method, which applies a correction factor to the *V*$_p$ based on available data range [38]; (ii) from the volume of correlation as described here [39]; (iii) using machine learning method finding nearest structures in shape and size from PDB database (Shape&Size, implemented in DATCLASS Fortran binary as a part of ATSAS) (accessed 2 July 2022) [40]; (iv) from Bayesian inference with *M*$_w$ calculated from methods mentioned above as the evidence to estimate most likely *M*$_w$ (implemented in DATMW program as a part of ATSAS) (accessed 2 July 2022) [41]. Normalized Kratky, Porod-Debye, and Kratky-Debye plots and volumetric analysis (calculation of Porod exponent, *V*$_p$, and molecular density) were prepared in ScÅtter 4.0 program (accessed 2 July 2022) (Harwell Science and Innovation Campus, Didcot, UK).

2.10. SAXS Data Processing

Electron density reconstruction was performed using DENSS program [42] implemented in BioXTAS RAW (accessed 2 July 2022), using previously computed *P*(*r*) as the only input, in slow mode and without any pre-determined symmetry with a number of electrons = 10,000. For averaging and refinement, at least 19 different reconstructions were aligned to obtain each averaged and refined reconstruction. Atomic representations of chimeric proteins were built from existing PDB structures (1DDT, 3LB6 for DT390-IL-13

and 1IKQ, 3LB6 for Exo) in UCSF Chimera (version 1.14) (accessed 2 July 2022) [43]. Missing loops were added using Modeller (accessed 2 July 2022) [44], remaining linkers and additional residues were constructed from Dunbrack2010 rotamers library [45], and such modified structures were merged to create complete atomic models. Local stereochemistry of these models was optimized by molecular dynamics using locPREFMD method [46] implemented in webserver (http://feig.bch.msu.edu/locprefmd/ accessed 2 July 2022) to remove steric clashes and improper rotamers and to add missing hydrogen atoms. Final optimized atomic structures were then used for further studies. Methods of ensemble generations: (i) *BILBOMB*48 implemented in the web server (https://bl1231.als.lbl.gov/bilbomd, accessed 2 July 2022) using 800 conformations per R_g value, R_g range ± 4 from experimental R_g determined from $P(r)$ function, flexible fragments defined manually; (ii) *Multi-FoXS*49 implemented in the webserver (https://modbase.compbio.ucsf.edu/multifoxs/ accessed 2 July 2022) using 10,000 conformations with manually defined rigid domains and without imposed distance constraints; (iii) *SREFLEX*50 implemented in *ATSAS online* web server (https://www.embl-hamburg.de/biosaxs/atsas-online/sreflex.php accessed 2 July 2022) in automatic mode using CONCORD (accessed 2 July 2022) during second refinement stage. All structural alignments of conformers and other manipulations or preparation of .pdb files were performed in the ChimeraX program (version 1.2.5) (accessed 2 July 2022) [47]. For each ensemble, a $P(r)$ function was calculated using BIFT, and these $P(r)$ functions were compared to appropriate experimental $P(r)$ functions to find ensembles best fitting to experimental SAXS data. Selected conformers generated by MultiFoXS (accessed 2 July 2022) were subjected to rigid docking in FoXSDock web server (https://modbase.compbio.ucsf.edu/foxsdock/) (accessed 2 July 2022) [48] using default parameters. For the selected dimeric complexes from FoXS-Dock (accessed 2 July 2022), $I(q)$ profiles were calculated using the FoXS web server (https://modbase.compbio.ucsf.edu/foxs/) (accessed 2 July 2022) [49], and corresponding $P(r)$ functions were calculated using BIFT and compared to experimental $P(r)$ function of the covalent dimer. Monomeric and dimeric atomic models were manually superimposed to ED reconstructions in ChimeraX (accessed 2 July 2022), and volumes of ED maps were adjusted to values estimated from the earlier volumetric analysis. Proposed trimeric and oligomeric structures were manually prepared in ChimeraX (accessed 2 July 2022) using dimers from FoXSDock (accessed 2 July 2022) and superimposed to ED maps, their $I(q)$ profiles were calculated using FoXS (accessed 2 July 2022) webserver and $P(r)$ functions were calculated by BIFT and compared to corresponding experimental $P(r)$ functions. All visualization of protein structures and ED maps were prepared in ChimeraX (accessed 2 July 2022).

2.11. Disulfide Bonds Mapping and Structural Analysis

Monomeric and oligomeric fractions were analyzed on a tandem LC-MS/MS system composed of the nanoAcquity UPLC (Waters Corporation, Milford, MA, USA) coupled to the Q Exactive Hybrid Quadrupole-Orbitrap mass spectrometer (Thermo Fisher Scientific, Waltham, MA, USA). Protein samples were non-reductively trypsin-proteolyzed for 10 h, acidified with trifluoroacetic acid (to 0.1%), and resolved by C-18 reverse-phase liquid chromatography on the trapping column (M-Class Symmetry C18 Trap Column, 100 Å, 5 μm, 180 μm × 20 mm) (Waters Corporation, Milford, MA, USA) and subsequently transferred online and resolved on the nano-HPLCRP-18 separation column (M-Class Peptide BEH C18 Column, 130 Å, 1.7 μm, 75 μm × 250 mm) (Waters Corporation, Milford, MA, USA) using a linear gradient of acetonitrile with 0.1% TFA (0–35% per 160 min, 0.25 mL/min flow rate and T = 40 °C). Data collection parameters were: data-dependent acquisition (DDA) and Top12 mode, scan range 300–2000 m/z, HCD fragmentation, mass resolving power 70,000 for MS1 and 35,000 for MS2, maximum ion accumulation time 60 ms for MS1 and 300 ms for MS2, automatic gain control (AGC) target value 10^6 for MS1 and $2 \cdot 10^5$ for MS2, isolation window 3.0 m/z, normalized collision energy (NCE) 27, and dynamic exclusion of 30 s. Raw data processing, peptide and crosslink identification,

and hit assembly were accomplished in MaxQuant 2.0.3.1 (accessed 2 July 2022) [50] and the Andromeda search engine [51]. Intensity is calculated as the summed up eXtracted Ion Current (XIC) of all isotopic clusters associated with the identified peptide sequence. The quality of recorded peptide spectra matches was ranked according to the binomial-probabilistic Andromeda scoring model and posterior error probability (PEP) with 5% threshold. The target-decoy-based 1% false discovery rate (FDR) cutoff thresholds for a peptide, a modified peptide, and a protein were applied. Prediction of aggregation propensity was computed using Aggrescan3D 2.0 method [52] on the webserver (http://biocomp.chem.uw.edu.pl/A3D2) (accessed 2 July 2022) in dynamic mode and 10 Å as a distance for aggregation analysis. Using default parameters, flexibility analysis was performed using CABS-flex 2.0 method [53] on the webserver (http://biocomp.chem.uw.edu.pl/CABSflex2) (accessed 2 July 2022). The calculated conformational ensemble was manually docked to ED reconstruction in ChimeraX (accessed 2 July 2022).

2.12. Titration Enzyme-Linked Immunosorbent Assay (ELISA)

Titration ELISA with the extracellular domain of IL-13-RA2 human recombinant receptor protein (ACROBiosystems, Newark, DE, USA) was performed at room temperature (RT). Pierce Nickel Coated Plates, Clear, 8-Well Strips (Thermo Fisher Scientific, Waltham, MA, USA), pre-blocked with BSA, were incubated with 100 μL of the 1 μg·mL^{-1} receptor in Dulbecco's PBS (DPBS) (Biowest, Nuaille, France) for 1 h. Strip wells were subsequently washed with DPBS and blocked with 2% skimmed milk in DPBS overnight, at 4 °C. Different fractions of the cytotoxin ligands, monomer, dimer, and lower oligomers, were added to the wells in serial dilutions in the concentration range 100–0.0001 μg·mL^{-1} for monomer and dimer and 50–0.00005 μg·mL^{-1} for lower oligomers, all in the presence of 0.1% BSA, each in 6 replicates and incubated for 2 h. Next, the strips were washed and incubated with the primary antibody (working dilution 1:165) for 2 h and the secondary (working dilution 1:3000) antibody for 1 h. The primary antibody was rabbit polyclonal anti-IL-13 antibody 0.1 mg/mL (OriGene, Rockville, MD, USA) and secondary—goat anti-rabbit IgG Fc, HRP conjugated, polyclonal (Agrisera AB, Vännäs, Sweden). After thorough washing, the strips were developed with 0.222 mg·mL^{-1} ABTS in 50 mM sodium citrate pH 4.0, supplemented with 0.05% H_2O_2. Then, 405 nm absorbance data were collected in the Tecan Infinite M200 PRO plate reader and the Magellan *7.2* software (accessed 2 July 2022) (Tecan, Mannedorf, Switzerland) and fitted to the sigmoidal four-parameter dose-response model in the GraphPad Prism 9.3.1 software (GraphPad Software, San Diego, CA, USA). The following parameters were derived: bottom, top, Hill coefficient, and LogEC$_{50}$ with 95% confidence interval and adjusted R2 for monitoring the weighted (1/Y) non-linear least-square regression quality. The null hypothesis was examined with the extra-sum-of-squares F test. The ordinary one-way ANOVA performed the further multiple pairwise comparisons of the LogEC$_{50}$ parameter with Tukey's correction and reporting *p*-values. The data sets were also fitted to the quadratic regression model to derive the apparent dissociation constants (see [54] for details).

2.13. In Vitro Functional Assay

Human GBM cell lines U-251 MG and LN229 were obtained from Wake Forest Baptist Medical Center Comprehensive Cancer Center and authenticated with STR analysis by ATCC. Cultures of U-251 (overproducing IL-13RA2) or LN229 (very low expression of IL-13RA2) cells were maintained in DMEM HG medium supplemented with 10% (*v/v*) fetal bovine serum and antibiotics: penicillin/streptomycin (1% *v/v*) and gentamycin (0.06% *v/v*). Cells were grown at 37 °C in a humidified atmosphere with 5% CO_2. Cell viability was determined by the MTS method. Then, 1×10^3 cells were plated per well into a 96-well plate in quadruplicates for each tested concentration and allowed to attach overnight. Next, the medium was replaced with fresh medium supplemented with HSA (final concentration: 0.1% *v/v*) and 0.1, 1, or 10 ng/mL of IL-13-based cytotoxins (monomer, dimer, and oligomers) for 72 h. Then, 20 mL of MTS solution was added to each well. After

4 h of incubation, absorbance was measured at 490 nm (with reference wavelength 670 nm) in a TECAN Infinite M200Pro microplate reader (Tecan, Mannedorf, Switzerland). Data were obtained from three independent experiments and presented as a percentage of the control. Data were analyzed using GraphPad Prism 8 software (GraphPad Software, San Diego, CA, USA). Two-way ANOVA followed by Tukey's multiple comparison test was used to calculate the *p*-value and determine the statistical significance of the difference in the measured variables between the control and tested groups. A difference between control and tested groups was recognized as significant when $p < 0.05$.

3. Results and Discussion

3.1. Preparation and Purification of the Cytotoxin

The fusion cytotoxin was expressed in *E. coli* without an affinity tag, in the process consisting of three major steps. After a short induction (2 h), the recombinant protein was recovered from inclusion bodies and solubilized in guanidium hydrochloride and reducing environment. The solubilized protein was subjected to refolding by the rapid dilution method and subsequent dialysis. The refolded and soluble fraction was separated and purified by anion-exchange chromatography (IEX) followed by size-exclusion chromatography (SEC). The production yield was 0.6–1.9 mg·L^{-1} at 90–99% purity, which was determined by densitometry. Later, the fermentation was switched from shaker flasks and orbital shaker to the medium size bioreactor unit, increasing yield up to 6.1–6.9 mg·mL^{-1} and achieving high reproducibility. Pressure homogenization was also incorporated into the purification protocol, and the cytotoxin recovery was in the range of 3.3–6.8% for monomeric cytotoxin. However, we found that a noticeable amount of the protein forms oligomers and aggregates during preparation. Since it reduces the yield and purity of the final product, we decided to characterize their structural and functional properties.

Figure 1 depicts the results of IEX and SEC chromatography. SDS-PAGE resolved the obtained fractions. We found that the main constituent of all fractions is the recombinant cytotoxin, and both IEX and SEC separate different soluble oligomeric states of the same protein. The monomeric cytotoxin was eluted at 15% buffer B during IEX and at SEC elution volume (V_{SEC}) of 14.4 mL, whereas for the dimer, these parameters were 20% and 12.2 mL, respectively. The oligomers eluted with IEX ionic strength above 20%, usually 25% and 30%, and V_{SEC} between 8 and 12 mL (peak maximum at 9.5 mL). The higher the oligomers, the lower the V_{SEC} with aggregates eluting closely to the exclusion limit of the column. These higher molecular forms are easily disrupted in reducing conditions into monomers, which indicates that oligomerization is cystine dependent, and oligomers are probably formed during refolding step.

Assuming that both dimer and oligomers are built by native and physiologically active monomers, we undertook some attempts to recover monomeric protein. Firstly, we tested a straightforward reduction using 50–100 tris(2-carboxyethyl)phosphine (TCEP)-to-monomer molar ratio, which was found to be sufficient to disrupt oligomers. However, the reduction led to excessive aggregation, making this strategy ineffective (Figure 1B). The second strategy was to use oligomers as potential recyclable substrates for the solubilization step of our biomanufacturing process (Figure 1B). The chromatographic and SDS-PAGE gel profiles are similar to those obtained routinely, with a total yield of 6.8% (pure monomer) and some contamination by dimer and oligomers. Expression and purification of Exo protein were performed using the analogical protocol (see Figure S2 for details, Supplementary Materials).
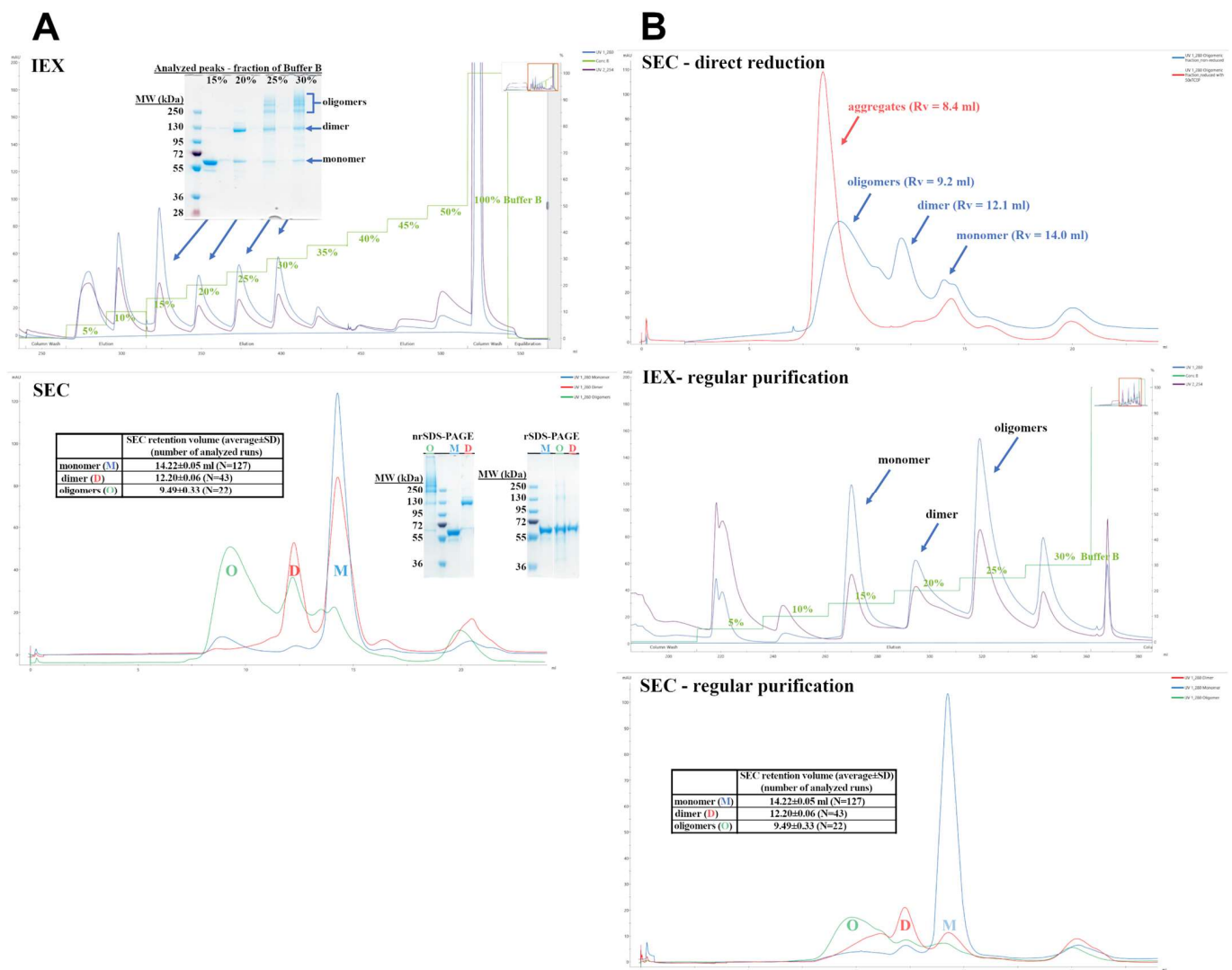
Figure 1. IEX and SEC chromatography of the cytotoxin. (**A**). IEX chromatogram with the recorded absorbance (254 nm, purple; 280 nm, blue) and conductivity (green) traces presents the initial separation of the protein dialysate into monomer, dimer, and oligomers. The peak fractions were resolved by non-reducing SDS-PAGE, revealing electrophoretic mobility of the monomer and dimer protein bands and a ladder of bands for the oligomer fractions. SEC chromatograms of the separated IEX peak fractions with the overlaid 280 nm traces are presented for the monomer (blue), dimer (red), and oligomers (green). The table inlet contains average elution volumes for each molecular state. The apex fractions were resolved by non-reducing and reducing SDS-PAGE. (**B**). The cytotoxin recovery from the oligomeric fractions by two tested approaches, direct reduction and regular purification protocol with IEX and SEC chromatography. Analytical SEC of the reduction substrates (blue) and products (red)—shifting the elution volume (V_{SEC}) for the monomer, dimer, and oligomers toward column void volume (about 8.0 mL) indicates aggregation. IEX chromatogram of the solubilized and refolded oligomers, separated into monomers, dimers, and oligomers. SEC chromatograms of the separated IEX peak fractions with the overlaid 280 nm traces are presented for the monomer (blue), dimer (red), and oligomers (green). The table inlet contains average elution volumes for each molecular state. X axes indicate elution volume (mL).

Our purification protocol is relatively simple and costs less than eukaryotic expression systems, and the recombinant protein requires no tag removal. Nevertheless, the production yield is relatively low (~10%), and the protein is prone to oligomerization. Although the oligomers can be resolved by a combination of IEX and SEC, their formation results in

monomeric product depletion, lowering the method's efficiency. Noticeably such behavior is a well-recognized phenomenon in the manufacturing of therapeutic antibodies [55]. However, it has not been reported for fusion cytotoxins yet.

Although the utilization of many novel expression systems for biomanufacturing biotherapeutics is constantly growing [56], protein expression in *E. coli* is still most prevalent in both industry and academia, also being usually the first-choice method to test protein production. Protein expression into inclusion bodies has long been distinguished as facilitating tag-free and protease-resistant protein manufacturing. More than 20 recombinant proteins are manufactured by recovery from inclusion bodies and refolding into soluble active products for industrial and biomedical applications [57].

3.2. Stoichiometry of the Oligomers

Dimers and low mass oligomers were resolved by non-reducing SDS-PAGE, whereas only a single band representing monomeric form was visible on reducing SDS-PAGE. This straightforward assay confirmed that oligomerization is cystine-dependent and assessed their electrophoretic mobilities (R_f) and stoichiometries (Figure 2).
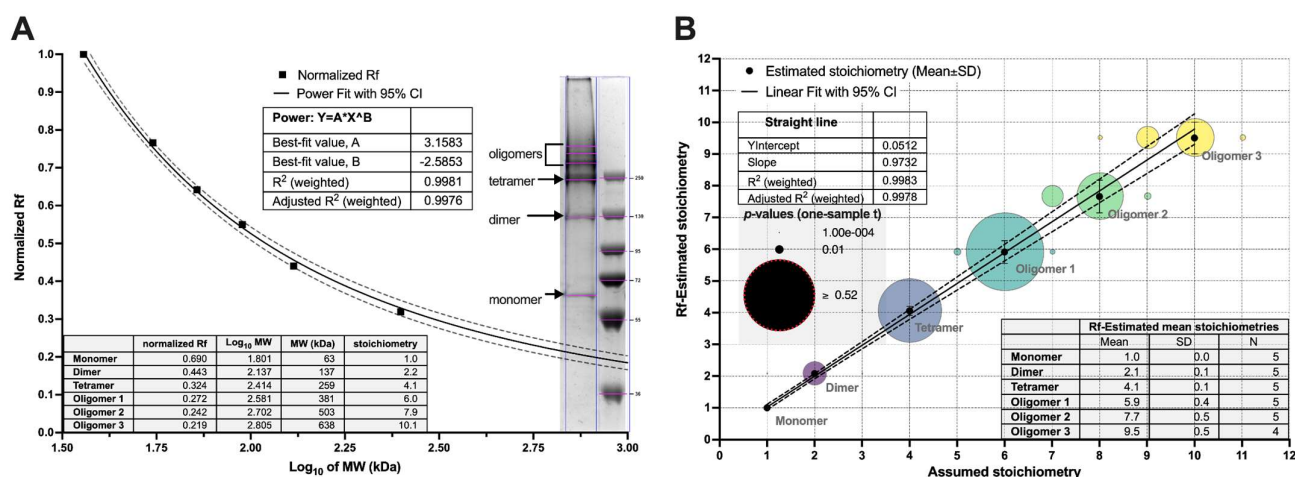

Figure 2. Determination of oligomers stoichiometry in non-reducing SDS-PAGE gels by analyzing relative front parameter (R_f). (**A**). Exemplary non-linear regression of the standard molecular weights plotted against their normalized relative fronts. The power model parameters were used to derive an apparent molecular weight of the cytotoxin bands, and their stoichiometries averaged for several gel paths. (**B**). The linear correlation between R_f-estimated and assumed stoichiometries for the dimer and oligomers. The R_f value for each band was also compared to the theoretical stoichiometries by the one-sample *t*-test, and the resulting *p*-values are displayed on the bubble plot. Each band was compared to three or four most probable values. Dotted lines represent the 95% confidence intervals (CI) for the regression models.

The first two oligomeric bands were in the range of 250 kDa. Their apparent molecular masses (M_{W-app}) were estimated by fitting a power function as a function of R_f. Their values were 62.7 ± 3.1, 123.1 ± 6.4, and 243.4 ± 6.1 kDa, indicating the following stoichiometries: monomer, dimer, and tetramer, respectively. To estimate stoichiometries of other oligomers, we applied analysis of their relative fronts (Figure 2), which resulted in the following numbers: 2.1 ± 0.1, 4.1 ± 0.1, 5.9 ± 0.4, 7.7 ± 0.5, and 9.5 ± 0.5. These were compared with hypothetical stoichiometries by the one-sample *t*-test (Figure 2). Although the most statistical significance may be attributed to the envisioned stoichiometries: 2, 4, 6, 8, and 10, more uncertainty was detected for octamer and decamer because of the reduced resolving power of SDS-PAGE. Moreover, dimers and tetramers are prevalent forms of DT390-IL-13 complexes, and the higher oligomers are relatively scarce. The obtained results suggest that the cytotoxic complexes are formed by dimeric units, which was further confirmed by the results of MALS and SAXS.

3.3. Computational Structural Studies

Comparative structural studies on our fusion cytotoxin in comparison to the native diphtheria toxin (DT) were analyzed by molecular dynamics (MD) simulations (Figure 3 and Table S1, Supplementary Materials). Five hypothetical models of DT390-IL-13 and a model of DT based on X-ray structures were subjected to 20 ns simulations.
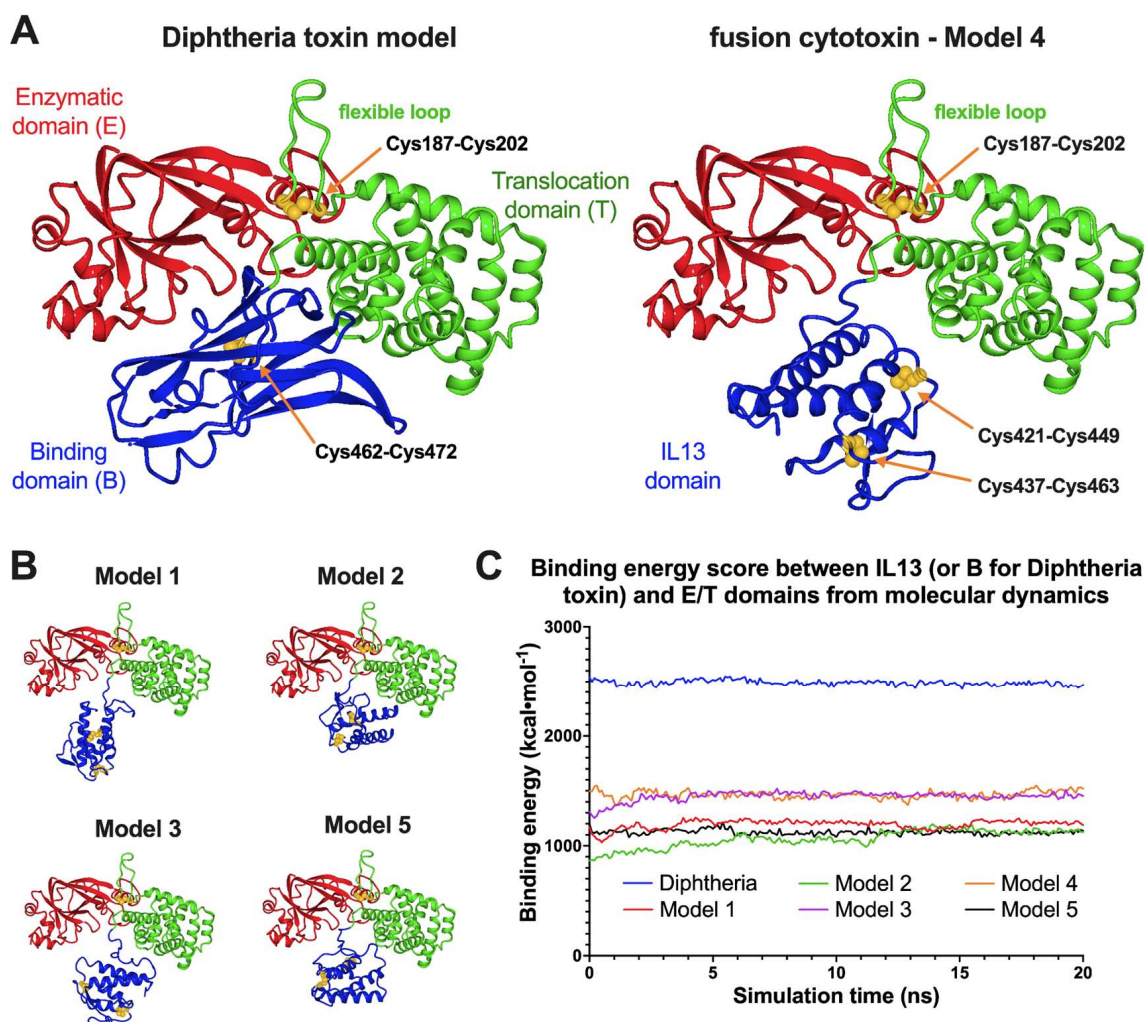
Figure 3. Molecular dynamics simulations for the diphtheria toxin and the fusion cytotoxin monomer. Presented are the ribbon representations of the last snapshots of the 20 ns MD. (**A**). Diphtheria toxin model compared to the cytotoxin Model 4, the one with the highest average DT390 and IL-13 interdomain binding energy score. (**B**). Structures of four other constructed models. (**C**). Trajectories of DT and DT390-IL-13 models as "binding energy" parameter in the function of time.

Values of time-averaged binding energies between IL-13 or the DT receptor binding domain and the enzymatic/translocation domains were significantly lower for the recombinant cytotoxin than for DT (2489 kcal·moL^{-1}), with the highest for Model 4 (1465 kcal·moL^{-1}) (Table S1, Supplementary Materials). Since the DT model was directly based on the low-energy crystallographic structure, it is expected to be more stable than the fusion protein with non-optimal interactions between both domains. The computed volumetric parameters ($\overline{V}$, R_0) are based on the single amino acid properties, so they are identical for all the models. Other parameters, including R_g, D_{max}, a/b, f/f_0, and *phi*$_c$ for both proteins, are similar and confirm their average prolate geometry (see Table S1 for the details). The slight differences probably were caused by replacing the binding domain with IL-13 in the DT390-IL-13 fusion protein. The values of R_g and R_t for the monomeric cyto-
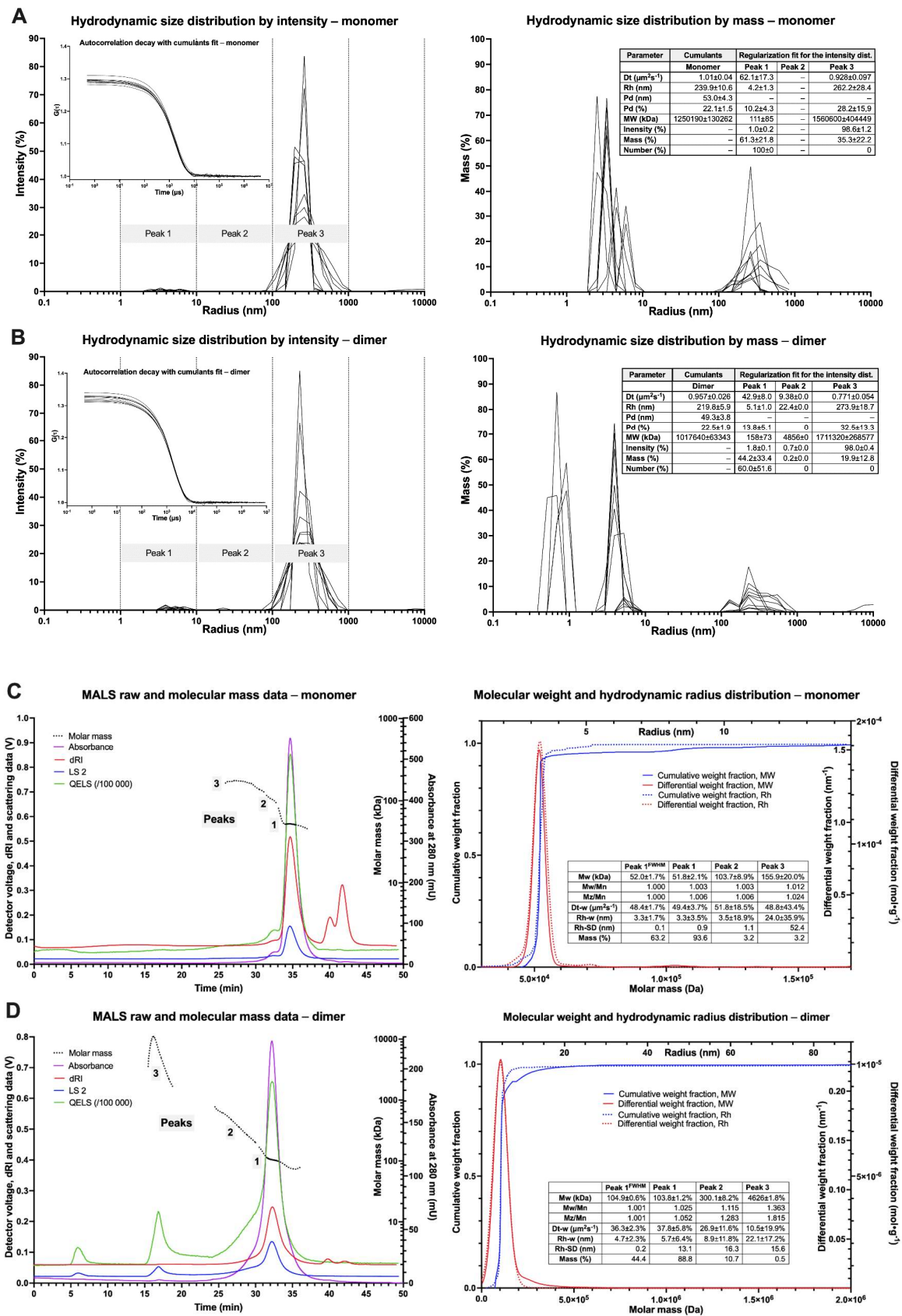
toxin models are in the ranges of 26.2 to 28.1 Å and 33.3 to 34.5 Å, respectively, whereas for DT, their values are 25.4 Å and 32.8 Å, respectively. The highest values of shape parameters were calculated for the "extended" Model 1 and the lowest—for the most energetically stable Model 4. The asphericity coefficient is higher for DT (0.21) than for the cytotoxin models (0.15–0.17). It can be concluded that swapping the receptor-binding domain for the IL-13 domain in the cytotoxin monomer may have a negligible influence on the molecule's size and shape. Nevertheless, the recombinant cytotoxin is expected to possess a less compact structure than the native toxin, due to the lack of specific intramolecular interaction between both domains.

3.4. Light Scattering Spectroscopy—DLS

Hydrodynamic properties, including R_h, dispersity, and modality, were determined by the batch DLS measurements using BSA as a reference (Figures 4 and S3). BSA sample is characterized by narrow and consistent intensity- and mass-averaged size distributions with a negligible mass fraction (about 0.3%) of particles with R_h above 10 nm. The R_h value for BSA was 4.1 ± 0.1 or 3.8 ± 0.1 nm for cumulant and regularization methods, respectively, which indicates a monomodal sample with a low polydispersity (Figure S3). On the contrary, for the monomeric fraction, large particles (>100 nm) were predominant, whereas small ones (<10 nm) contributed only about 1.0%. This sample demonstrated much higher cumulant R_h (239.9 ± 10.6 nm) and its polydispersity ($53.0 \pm 4.3\%$) but similar R_h values for Peak 1 (4.2 ± 1.3 nm), which corresponds to monomeric particles. Comparable results were obtained for the dimer sample, where the Peak 1 radius was slightly higher (5.1 ± 1.0 nm). Nonetheless, both samples are statistically indistinguishable by DLS. These results suggest a tendency of the cytotoxin molecules to aggregate and oligomerize in the solution. In the case of the oligomer fraction, the particles with radii in the range of 5–50 nm were responsible for 93.0% of the intensity. Mass-averaged distribution for the oligomers revealed that about half the mass of scattering particles come from smaller scatterers with the $R_h = 2.4 \pm 0.9$ nm and another half from larger scatterers with $R_h = 14.9 \pm 1.5$ nm suggesting bimodal distribution. On the other hand, the characteristic of Peak 2 is similar to cumulant data ($R_h = 12.6 \pm 0.4$ nm, polydispersity 3.0 ± 0.1 nm), suggesting monomodality with the predominant oligomeric particles (Peak 2). These findings suggest a general tendency for oligomerization and aggregations of DT390-IL-13 with high-mass particles in both the monomer and dimer samples.

### 3.5. Light Scattering Spectroscopy—MALS

To reduce the ambiguity of the results of batch DLS, we applied SEC-MALS-DLS to analyze the size distribution of DT390-IL-13 (Figure 4C,D) from different fractions. The separation of BSA, used as a control, resulted in peaks corresponding to monomer (95.1%), oligomers (4.1%), and aggregates (0.8%). The cytotoxin monomer sample was resolved into three peaks, monomeric (93.6%) and two oligomeric (6.4%). The measured M_w for the monomer was 52.0 kDa showing high similarity to the actual M_w (55.1 kDa). The R_h value determined by MALS (3.3 ± 0.1 nm) was more precise than the DLS results. For the dimeric sample, three MALS peaks were found, corresponding to the following molecules: dimer (88.8%), oligomers (10.7%), and aggregates (0.5%), with M_w and R_h for the dimeric peak of 104.9 kDa and 4.7 ± 0.2 nm, respectively. The dimer-to-monomer M_w ratio was 2.0. Dispersity parameters, M_w/M_n and M_z/M_n were virtually the same for the dimeric fraction (1.001) in comparison to the monomeric fraction and BSA (1.000). These parameters were higher for the full peaks of the monomer and dimer than BSA, which suggests slight amounts of oligomers. In the case of the oligomeric sample, four different scatterers sets were found a (M_w 258.5 kDa, R_h, 7.9 + 0.5 nm, t_{el} 32 min), b (high M_w, R_h 8.9 + 0.7 nm, t_{el} 25–30 min), c (high M_w, R_h 11.8 + 1.2 nm, t_{el} 20–25 min) and d (M_w beyond the detection limit), which indicate high polydispersity (Figure S3, Supplementary Materials). Sets b and c had the highest mass contribution (45.0% and 34.2%, respectively) and their M_w was more sparsely distributed when compared to a.

A

Hydrodynamic size distribution by intensity – monomer

Autocorrelation decay with cumulants fit – monomer

Peak 1   Peak 2   Peak 3

Hydrodynamic size distribution by mass – monomer

Parameter	Cumulants	Regularization fit for the intensity dist.		
	Monomer	Peak 1	Peak 2	Peak 3
Dt ($\mu m^2 s^{-1}$)	1.01±0.04	62.1±17.3	–	0.928±0.097
Rh (nm)	239.9±10.6	4.2±1.3	–	262.2±28.4
Pd (nm)	53.0±4.3	–	–	–
Pd (%)	22.1±1.5	10.2±4.3	–	28.2±15.9
MW (kDa)	1250190±130262	111±85	–	1560600±404449
Intensity (%)		1.0±0.2	–	98.6±1.2
Mass (%)	–	61.3±21.8	–	35.3±22.2
Number (%)	–	100±0	–	0

B

Hydrodynamic size distribution by intensity – dimer

Autocorrelation decay with cumulants fit – dimer

Peak 1   Peak 2   Peak 3

Hydrodynamic size distribution by mass – dimer

Parameter	Cumulants	Regularization fit for the intensity dist.		
	Dimer	Peak 1	Peak 2	Peak 3
Dt ($\mu m^2 s^{-1}$)	0.957±0.026	42.9±8.0	9.38±0.0	0.771±0.054
Rh (nm)	219.8±5.9	5.1±1.0	22.4±0.0	273.9±18.7
Pd (nm)	49.3±3.8	–	–	–
Pd (%)	22.5±1.9	13.8±5.1	0	32.5±13.3
MW (kDa)	1017640±63343	158±73	4856±0	1711320±268577
Inensity (%)	–	1.8±0.1	0.7±0.0	98.0±0.4
Mass (%)	–	44.2±33.4	0.2±0.0	19.9±12.8
Number (%)	–	60.0±51.6	0	0

C

MALS raw and molecular mass data – monomer

Molar mass
Absorbance
dRI
LS 2
QELS (/100 000)

Peaks   3   2   1

Molecular weight and hydrodynamic radius distribution – monomer

Cumulative weight fraction, MW
Differential weight fraction, MW
Cumulative weight fraction, Rh
Differential weight fraction, Rh

	Peak 1^FWHM	Peak 1	Peak 2	Peak 3
Mw (kDa)	52.0±1.7%	51.8±2.1%	103.7±8.9%	155.9±20.0%
Mw/Mn	1.000	1.003	1.003	1.012
Mz/Mn	1.000	1.006	1.006	1.024
Dt-w ($\mu m^2 s^{-1}$)	48.4±1.7%	49.4±3.7%	51.8±18.5%	48.8±43.4%
Rh-w (nm)	3.3±1.7%	3.3±3.5%	3.5±18.9%	24.0±35.9%
Rh-SD (nm)	0.1	0.9	1.1	52.4
Mass (%)	63.2	93.6	3.2	3.2

D

MALS raw and molecular mass data – dimer

Molar mass
Absorbance
dRI
LS 2
QELS (/100 000)

Peaks   3   2   1

Molecular weight and hydrodynamic radius distribution – dimer

Cumulative weight fraction, MW
Differential weight fraction, MW
Cumulative weight fraction, Rh
Differential weight fraction, Rh

	Peak 1^FWHM	Peak 1	Peak 2	Peak 3
Mw (kDa)	104.9±0.6%	103.8±1.2%	300.1±8.2%	4626±1.8%
Mw/Mn	1.001	1.025	1.115	1.363
Mz/Mn	1.001	1.052	1.283	1.815
Dt-w ($\mu m^2 s^{-1}$)	36.3±2.3%	37.8±5.8%	26.9±11.6%	10.5±19.9%
Rh-w (nm)	4.7±2.3%	5.7±6.4%	8.9±11.8%	22.1±17.2%
Rh-SD (nm)	0.2	13.1	16.3	15.6
Mass (%)	44.4	88.8	10.7	0.5

Figure 4. Results of DLS and MALS measurements for monomeric and dimeric fractions of DT390-IL-13. (**A**) DLS of monomeric fractions. (**B**) DLS of dimeric fraction. (**C**) MALS of monomeric fractions. (**D**) MALS of dimeric fraction.

Separated oligomeric fractions were subjected to nrSDS-PAGE, which resolved monomer, dimer, and oligomers (Figure 5). As expected, oligomers M_w was inversely proportional to t_{el}. The oligomers, as well as aggregates, were easily monomerized after reduction by DTT. Protein quantity is decreasing for the fractions of high oligomers to the identified Peak 3 on SEC-MALS (Figure S3, Supplementary Materials), and the last three fractions are devoid of the protein. This indicates that particles present in this peak are unstable, and experimental manipulations accelerate their denaturation and precipitation. The M_w values determined by MALS and similar methods are weight-averaged (Equation (1)), and the accurate distribution of different particles in the polydisperse system may be challenging to estimate without additional experimental evidence.

$$\overline{M_w} = \frac{\sum c_i M_i}{\sum c_i} \tag{1}$$
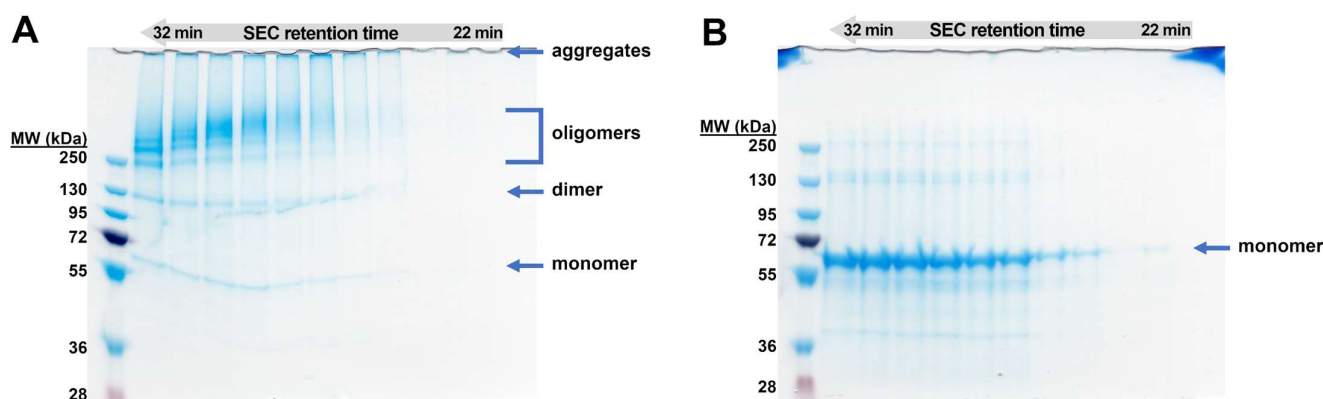
Figure 5. SDS-PAGE analysis in non-reducing (**A**) and reducing (**B**) conditions of the consecutive SEC-MALS oligomer fractions with the elution time range of 22–32 min. An amount of 10 μg was loaded into each well. The average mass distribution over the analyzed fractions in the non-reducing conditions is inversely proportional to the SEC-MALS elution time, as expected. The higher the molecular weights, the more unstable the cytotoxin. The oligomeric fractions additionally contain monomer, dimer, and aggregates. The efficient oligomers monomerization was observed in the reducing conditions.

SDS-PAGE analysis of the low-oligomeric fractions indicated that the mass fraction is inversely proportional to the oligomer stoichiometry. The oligomer fractions also contain some amount of the monomer and the dimer. This way, the MALS-determined M_w of 258.5 kDa could be potentially resolved into its ingredients, monomers, and oligomers with the $2n$ stoichiometry. Assuming the monomer and dimer contents are each less than 12–20% and the octamer/decamer are at the low level of 1–2%, the equation-calculated tetramer and hexamer fractions would be in the range of approximately 26–43%. These results are in accordance with DLS and were further confirmed by SAXS.

Oligomeric samples are polydisperse, and oligomerization generates rather spheroidal than rod-like particles. It is believed that most proteins are prone to (self)-assembly, especially in higher concentration and non-biological environments, and sometimes even a single point mutation may increase such tendencies. Our findings suggest that the oligomerization may not be governed only by a random disulfide rearrangement and misassemble but is also supported by non-specific intermolecular interactions in the vicinity of Cys residues involved in the formation of disulfide bonds. The notion may be supported by the fact that the Exo protein showed no oligomerization tendency (Figure S2), at least in the preliminary experiments, yet still may form aggregates. Thus, oligomerization might be triggered by the presence of solvent exposed Cys residues in the flexible region of DT390.

3.6. Small-Angle X-ray Scattering

3.6.1. Decomposition of SAXS-SEC Data

For further biophysical 1analysis, we applied SEC-SAXS to achieve better buffer subtraction, separate different molecular species, and remove aggregates. A set of samples containing different oligomeric states of DT390-IL-13 and Exo proteins was (datasets 1.0–5.0) ubjected to SEC-SAXS. Initial $I(q)$ profiles indicated flexibility and potential aggregation in each sample subjected to SEC-SAXS (Figure S4, Supplementary Materials).

Despite the application of SEC-SAXS to separate different oligomeric states of the proteins, initial data analysis, e.g., the shape of scattering profiles and $P(r)$ functions of datasets 1.0–5.0, indicated polydispersity and aggregation. In order to identify a number of scattering components and their relative contributions, Singular Value Decomposition (SVD) was used. SVD is a data reduction technique based on the factorization of data matrix A to extract its main components. SVD decomposes the scattering data into orthonormal components, known as singular vectors, ranked in order of significance by the singular values (SV). Detailed mathematical descriptions can be found elsewhere [58], and a brief description is available in ESI. Our SVD analysis indicated a mixture of overlapping scattering species in each fraction, rendering straightforward data processing complex since automatic methods were unable to discern sets of scattering profiles suitable for further processing (see Figure S5, Supplementary Materials). To overcome this hurdle, we applied Evolving Factor Analysis (EFA), a model-free data deconvolution method suitable for time-dependent datasets [35,59]. Using EFA, we resolved overlapping signals from all SEC-SAXS datasets, except 2.0, when it led to unphysical results (i.e., negative concentrations). In some cases, complete separation of different species was not possible due to large numbers of SV and relatively low signal-to-noise values indicated by high χ^2 error (Figure S6, Supplementary Materials).

3.6.2. Selection of Suitable Datasets

Scattering profiles for each SV were further analyzed to assess their suitability for further processing, and the results were compiled in Table S3, Supplementary Materials. Since both studied proteins are composed of two parts connected via a linker, it was expected that at least monomeric forms are flexible systems, which should be confirmed by analysis of SAXS data. To select suitable data sets, we used the following criteria: reasonable values of radius of gyration (R_g), M_w expected for a given fraction, e.g., monomer, dimer, and overall shape of $P(r)$ function (Figure 6).
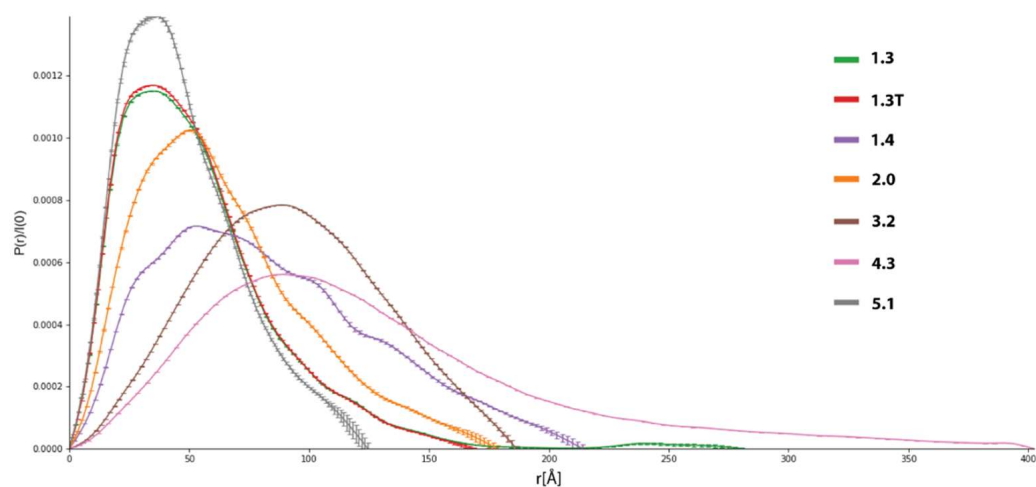
Figure 6. $P(r)$ functions for selected datasets extracted from initial SEC-SAXS data. All these datasets were selected for further analysis.

$P(r)$, also called the pair-distance distribution function, is a measure of the frequency of interatomic vector lengths within a scattering particle, providing basic information

about the shape of the particle. R_g values were obtained from either Guinier analysis or from the $P(r)$ function. For most datasets, R_g determined from $P(r)$ was slightly higher than Guinier's approximation, indicating conformational flexibility or the presence of oligomers. Substantial differences between these values were observed in some datasets, suggesting aggregation or denaturation. We calculate $P(r)$ functions using a fully automated procedure based on Bayesian Indirect Fourier Transform (BIFT), which always yields one $P(r)$ function of the highest probability. The Bayesian approach to the Indirect Fourier Transform (IFT) allows us to obtain an objective estimate of the $P(r)$ without any human interaction, significantly reducing 'qualified biasing' of the statistical parameters in terms of the maximum distance within the particles (D_{max}), and the smoothness constraint α. Hence, the BIFT algorithm provides a more objective estimation of $P(r)$ compared to standard IFT algorithms [43]. Datasets indicating severe or medium aggregation were excluded from further processing (Figure S6, Supplementary Materials). On the other hand, in each sample, at least one dataset displays well-behaved $P(r)$ functions, which was the vital criterion in this assessment. For each dataset, M_w was calculated using several different methods (based on Porod volume V_p, the volume of correlation V_c, Bayesian inference, and comparison to known structures, See Material and Methods for details), and for most datasets, they converged to similar values. In our case, the lowest estimates for M_w are in the 80–90 kDa range, while the real *M*ws are significantly lower (~52 kDa). M_w for flexible or disordered systems determined from $I(q)$ profiles tends to be overestimated due to higher apparent volume in the solution [60]. On that ground, we assumed that low-mass fractions contained mostly monomers, and the mass error was caused mainly by their dynamic behavior rather than contamination by larger molecules. It should also be emphasized that in SAXS and other scattering techniques, even a small contribution of massive particles leads to a noticeable overestimation of molecular volume, M_w, and shape parameters such as R_g or persistence length, also reflected in our data [60].

Considering the abovementioned criteria, we selected several datasets for more detailed studies. The analysis of SV for sample 1.0 showed its higher diversity than other samples subjected to SEC-SAXS, since it contained mono- and dimeric protein and oligomers and aggregates. Dataset 1.2 contains oligomers, and 1.3 (and its truncated version 1.3T)) is mainly composed of monomers and some contributions of dimers and oligomers, making it suitable for modeling. Dataset 1.4 contains mostly trimers/tetramers and was also selected for modeling. In the case of 2.0, only a basic background correction was performed, and a set of few scattering profiles was selected based on the shape of its $P(r)$ function, indicating the presence of only one flexible species. Fractions 3. 2 and 4.3 are composed of a mixture of oligomers and aggregates, and after EFA, one most promising elution profile from each fraction (3.2 and 4.1) was selected for modeling and ensemble analysis. Dataset 5.0 is the sample of Exo protein-containing aggregates, and an elution profile 5.1 corresponding to monomeric protein was extracted. Other datasets were discarded from further processing.

3.6.3. Further Analysis of SAXS Data for Selected Datasets

There are a number of indicators of flexibility and disorder derived from I(q) profiles, including numerical parameters such as Porod exponent d and the general shape of Porod-Debye (PD), Kratky, or Kratky-Debye plots [61]. In our assessment, we firstly analyzed the shape of PD and dimensionless Kratky plots (Figures S7 and S8, Supplementary Materials). Datasets of lower M_w (1.3, 1.3T, 1.4, 2.0, and 5.1) display partial flexibility, typical for multidomain proteins, while systems having high M_w (3.2 and 4.3) are more rigid. Shapes of PD plots also indicated that low M_w datasets contain flexible particles. A plateau in the PD plot is expected for globular systems, while it is not present for flexible or disordered molecules. The volumetric analysis allowed us to obtain d, Porod Volume (V_p), and particle density ρ (Table S4, Figure S9, Supplementary Materials). Lower values of d (<3) indicate flexible or disordered protein chains, while values above 4 indicate rigid particles. Molecular shapes estimated from d correspond well to ones estimated from Kratky and PD plots. V_p values for 1.3, 1. 3T, 1.5, and 2.0 are considerably lower, suggesting they are

composed mainly of monomers or dimers (in the case of 2.0), while others are oligomeric. Low values of ρ also suggest flexible systems since this parameter is determined from V_p, which is overestimated for such particles. Finally, integrated total scatter intensity plots over q have a slight slope at high q values, indicating a lack of severe aggregation or interparticle interferences [39]. Analysis of scattering data in real space via the) P(r) function also provides information about the shape and behavior of scattering particles. Here, these are in agreement with the reciprocal space analysis discussed earlier (Figure 6 and Figure S10, Supplementary Materials). Elongated tail for high D_{max} values indicates a contribution from oligomers or aggregates in the 1.3 dataset and, to a lesser degree, for 1.4, 2.0, and 5.1. All mentioned datasets possess right-skewed pair-distribution suggesting partial flexibility or the presence of two connected domains. The P(r) shape for the 3.2 dataset indicates a more rigid structure resembling a hard sphere. Dataset 4.2 possess elongated P(r) distribution skewed to very high D_{max} values due to significant aggregation. Moreover, it confirms that some datasets (notably 1.3 and 4.2) are contaminated by aggregates, which may hinder structural modeling. Hence, we recalculated P(r) for 1.3 datasets with additional restraint for very high (>150 Å) D_{max} values. As a result, we obtained truncated dataset 1.3T, in which the P(r) function is almost the same as the unrestrained one, except for a significantly shortened tail. Similar recalculation failed for 3.2, yielding unphysical P(r) functions; thus, we decided to use the unmodified function for molecular modeling.

### 3.6.4. Electron Density Reconstruction and Ensemble Analysis

For 3D shape reconstruction, we decided to use the DENsity from Solution Scattering (DENSS) method to obtain electron density (ED) of particles directly from 1D scattering data, avoiding many assumptions limiting the accuracy of other modeling algorithms [42]. Since such reconstructions are non-uniform in terms of ED distribution, DENSS is more suitable to work with flexible or intrinsically disordered proteins than bead models. Flexible regions manifest as regions of partial occupancy where average ED is reduced. Datasets 1.3T, 1,4, 2.0, 3.2, 4.3, and 5.1 were subjected to DENNS, and in each case, except 4.3, it yielded models of satisfying quality (Table S5, Figures S11–S13, Supplementary Materials). The overall resolution (FSC cut-off 0.5) is moderate (~40 Å for 1.3T and 5.1 or ~55 Å for 1.4 and 2.0) to low (~100 Å for 3.2), which is caused by the innate flexibility of the molecules as well as minor data contamination by scattering from high M_w particles, artificially increasing the D_{max} values. In the case of 4.3, the very high χ^2 score indicates the reconstruction is not valid, thereby not acceptable for analysis.

Several X-ray structures of bacterial toxin and interleukin-receptor complexes are available in the PDB database. Nonetheless, due to the internal flexibility of our chimeric proteins, their structures in solution may significantly vary from the available X-ray structures adapting a set of different conformations. In a traditional approach, high-resolution sampling of conformational space at the atomic level is achieved through molecular dynamics (MD) or Monte Carlo simulations. On the other hand, when working with lower resolution SAXS data, either a coarse-grained approach or simplified MD often dispenses an acceptable resolution for a significantly reduced computational cost. We employed several different methods of ensemble generation and compared the results. In the BILBOMD method, a fast MD simulation of interdomain linkers at a very high temperature samples conformation space and generates a set of conformers from which an $I(s)$ curve is calculated and compared to the experimental data [62]. Later, a genetic algorithm performs a minimal ensemble search (MES) which provides a few subsets of conformers best fitting to the experimental scattering profile. Another approach to address conformational heterogeneity is to probe the space of the φ and ψ dihedral angles for previously defined flexible fragments using the Rapidly Exploring Random Trees algorithm, which is implemented in the Multi-FOXS webserver [48]. When a vast set of conformers is created, a $I(s)$ curve is calculated for each structure, and multi-state models are validated against the scoring function to select the most promising models. Another method is SREFLEX, a hybrid approach applying Normal Mode Analysis (NMA) to generate a large set of atomic models and then refine the

conformation by comparing their scattering profile to the experimental data [63]. Contrary to previous methods, SREFLEX does not compute a true conformational ensemble, but a set of different conformers gives similar scattering profiles to the experimental one.

We prepared atomic models of our proteins from several PDB structures, which were merged, rebuilt, and further optimized using local MD. These steps were necessary since improper input structures may result in computational errors or distorted output multi-stage models. All ensemble methods gave similar results, which is reflexed in the comparison of the $P(r)$ function of computed multi-states to experimental $P(r)$ functions for both 1.3T and 5.1 datasets (Figure S14, Table S6, Supplementary Materials). The results for the 1.3T dataset will be discussed first. BILBOMB generated several multi-state models consisting of 1 to 3 structures. For all ensembles, the $P(r)$ shapes suggest the presence of two separate regions (two distinct maxima of the probability distribution). The D_{max} values for each ensemble are much lower than for the experimental $P(r)$ curve since it does not take into account traces of oligomers presented in the data. Results obtained from SREFLEX are similar to BILBOMB, since the superposition of most relevant conformers resembles three-state models generated by the latter method. MultiFoXS generated similar multi-state models to these from BILBOMB, but D_{max} values are in better accordance with experimental results. IL-13 domain is located at a higher distance from DT390, which explains higher D_{max} values and suggests that the linker exists mainly in extended conformation providing maximal separation of the two regions. This is expected to facilitate intermolecular interaction between protein molecules leading to the formation of covalent dimers. Two-state and three-state MultiFoXS structural models were aligned to reconstruct ED from the 1.3T dataset (Figure 7A) at the density threshold necessary to recreate ED volume similar to V_p obtained from the volumetric analysis. For analysis of other ED reconstructions, we applied the same ED threshold. The superposition of the structural model and ED shows overall good structural alignment, in which DT390 fits into a bulkier fragment of ED.

In contrast, a smaller IL-13 domain occupies an oblong and thinner ED fragment. Despite the high flexibility of the interdomain linker, the ED coverage for IL-13 is still almost complete suggesting that statistical ensembles probed a significant part of the conformational space. Ensembles computed for the 5.1 dataset, representing monomer of Exo protein, indicate reduced conformational flexibility if compared to 1.3T (Figure S15). Both BILBOMB and MultiFoXS generated only single-state models, and models from SREFLEX display a high level of resemblance to one another. All methods provided comparable results, which are in good agreement with the experiment. The only exceptions are some models from BILBOMB and MultiFoXS, whose $P(r)$ deviates from the experimental data. Lack of either high-order molecular ensembles or conformational variety indicates reduced flexibility of the system compared to DT390-IL-13. These results also partially explain the lack of covalent dimers in the case of the Exo molecule, since the relative proximity of two domains impedes closer intermolecular interactions necessary to form covalent bonds. Without additional studies, it is difficult to explain the reduced flexibility of the Exo protein, since even the relatively short interdomain linkage provides significant conformation freedom. Structural alignment of the atomic model generated by MultiFoXS with reconstructed ED shows visible similarity indicating general agreement between these methods (Figure 7), with regions of higher electron density fitting considerably well to the main polypeptide chain (Figure S15, Supplementary Materials).
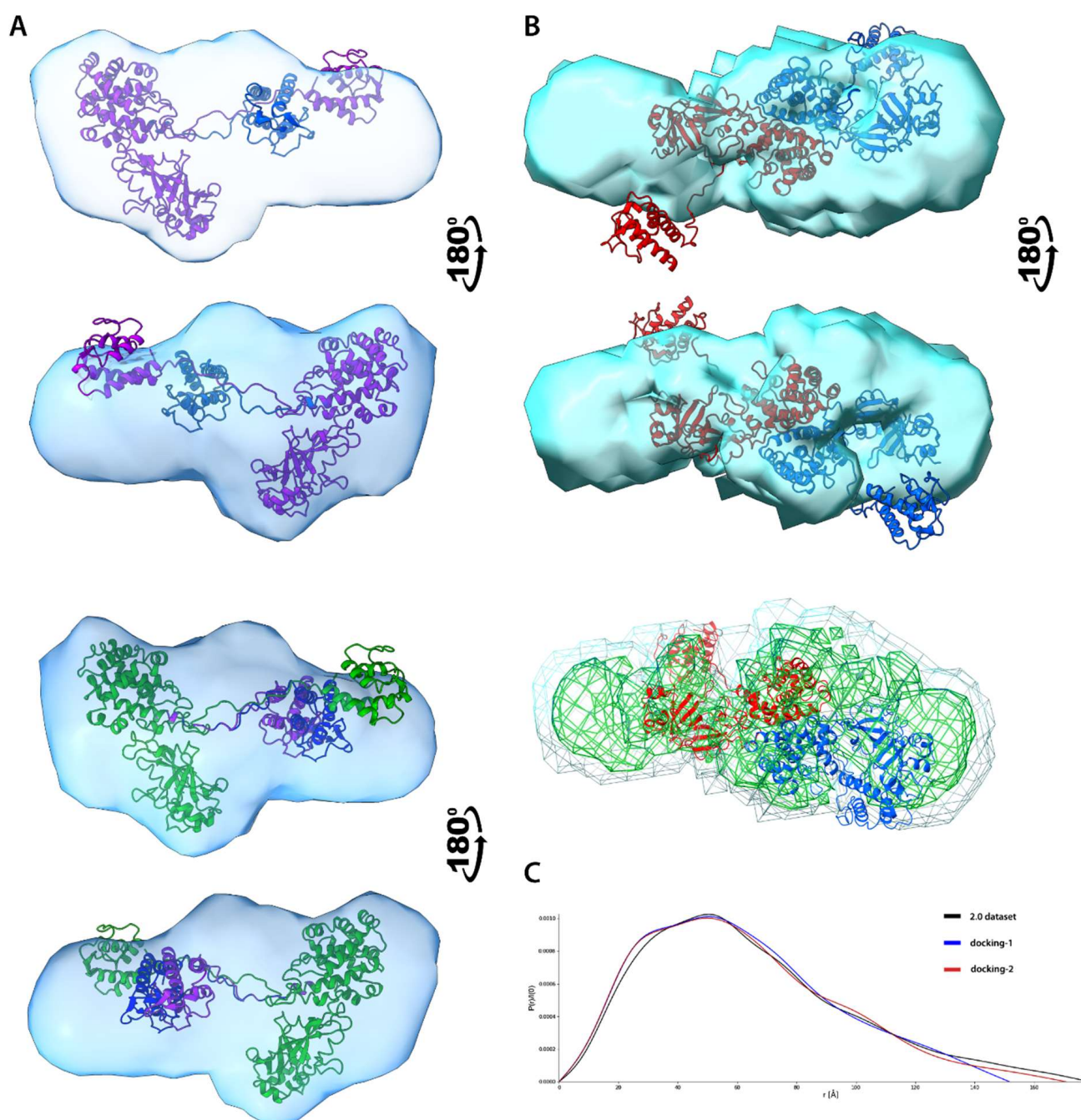
Figure 7. SAXS structural modelling of DT390-IL-13. (**A**). Alignment of atomic models representing 2-state (top) and 3-state (bottom) models onto ED reconstruction. Each conformer is depicted using a different color. (**B**). Alignment of atomic model of best-scoring dimer onto ED reconstruction, each monomer is depicted using a different color (top). Layers of different electron density showing coverage of polypeptide chain (bottom). (**C**). *P*(*r*) functions for both original and 2.0 dataset (representing a dimer)and two selected dimers created by the FoXSDock.

Structural characterisation of protein complexes using only SAXS data usually is difficult and poses a number of challenges, especially for flexible systems [60]. One method suitable for modelling dimeric complexes is webserver FoXSDock which allows for protein-protein rigid docking restrained by a scattering profile of the complex [48]. FoXSDock generates a list of dimeric complexes sorted by a combined SAXS and statistical potential energy score. We applied this tool to obtain estimated models of DT390-IL-13 dimers using selected models from previously generated statistical ensembles. Since MultiFoXS

ensembles were closest to experimental data we decided to select most prevalent conformers from this method, and following pairs of structures: *e1-1 + e1-1*; *e2-1 + e2-2*; *e3-2 + e3-2* and *e4-2 + e4-3* were subjected for modelling (see Table S6 for details, Supplementary Materials) using $I(s)$ profile of dataset 2.0. Since FoXSDock is unable to model covalent complexes, the possible output is only a rough approximation of the real covalent dimer and should be investigated with caution. To assess the quality of the dimers two criterions were taken into account: overall FoXSDock score and chemical plausibility giving the higher importance to the latter one. We discarded all complexes in which Cys residues were at large distances preventing them from formation disulfide bonds. Complexes possessing Cys residues in relative proximity allowing them to form covalent bonds after minor conformational adaptations such as repositioning of flexible loops or changes in their relative orientation were further evaluated on the base of their $P(r)$ functions (Figure S16). Among these complexes, it is possible to propose two different class of dimers which vary in their presumed topology of disulfide bonds: covalent bonds between two DT390 regions (called here *endo*-dimers) or bonds between DT390 and IL-13 (dubbed as *exo*-dimers). Noticeably, for one complex composed of two *e4-3* conformers, its $P(r)$ function is almost identical to the one computed from experiment (Figure 2C). It suggests that the real covalent dimer may possess very similar structure to the proposed endo-dimer depicted in Figure 7. The formation of disulfide bridges would require only a rotation (~30–40°) of one monomer in regard to its partner, to reduce the distance between Cys residues (Figure S17, Supplementary Materials) resulting in a better fit to the ED model. Fitting of the atomic model of that dimer into a reconstructed ED strengthen this claim. The overall structural alignment is fairly high, and the polypeptide chain of more rigid fragment comprising DT390 domains fits into a region of higher density (Figure 7B). On the other hand, the fitting of the IL-13 domain is worse, which is partly caused by its flexibility. If one DT390-IL-13 was rotated relative to the second molecule, as was postulated earlier, the fit of flexible parts would be considerably improved.

　　Structures of endo- and exo-dimers can be considered as composite blocks for covalent oligomers, which was already proposed on the ground of SDS-PAGE and MALS results. Due to steric restrains, any oligomeric structures are likely to be composed of molecules displaying both topologies of disulfide bonds forming an endo-endo-exo trimeric structural unit, a which structure can be approximated by the partial overlay of endo- and exo-dimers. Because of the lack of high-resolution structural data, we cannot exclude the existence of high-order complexes connected mainly via DT390 domains in an endo-endo fashion. However, it would be challenging to avoid significant structural clashes between relatively bulky and rigid DT390 domains in such a case. It is unlikely that any higher-order oligomers could be formed only via DT390 domains. Interestingly, the ED model for dataset 1.4 can be reasonably fitted to the approximate structure of such a trimer composed of molecules connected via DT390 and IL-13 (Figure S18, Supplementary Materials). Reconstruction of ED for a 4.3 dataset was only partially successful due to the limited resolution. It probably represents an ensemble of higher oligomers with a more rigid structure than monomers and dimers. Such compact structure requires a higher number of spatial restraints than flexible linear oligomeric chains, suggesting that higher oligomers may be cyclic or composed of partially misfolded protein chains forming globular particles via several non-specific interactions. These claims are supported by cytotoxic assays showing their limiting potency if compared to the monomer (See Paragraph 3.8 for details). On that ground, we may propose a hypothetical model of oligomerization in which a number of DT390-IL-13 molecules form caged structures connected by DT390 and IL-13 domains (Figure S19, Supplementary Materials). Nonetheless, it needs to be emphasized that such a model is highly speculative and requires further experimental validation due to the deviation of the experimental P(r) function from the theoretical one computed for our approximate structural model.

3.7. Disulfide Mapping and Structural Analysis

Peptide and disulfide crosslink mapping was accomplished by the bottom-up tandem MS with a semi-stochastic approach and HCD fragmentation. The overall sequence coverages were: 78% for the monomer, 68% for the dimer, and 71% for the oligomers. However, the highest-intensity peptides were detected for the enzymatic toxin domain and interleukin 13 domain (Figure 8 and Figure S20, Supplementary Materials). The former includes a disulfide bridge native to the toxin C187–C202. The following regions were poorly identified: the N-terminal part of the toxin translocation domain, the IL-13 peptide with two disulfide linkages, and the C-terminal amino acids. The long tryptic IL-13 peptide (residues 418–466) containing two native disulfide linkages, C421–C449 and C437–C463, is either poorly separated during LC or weakly fragmented.
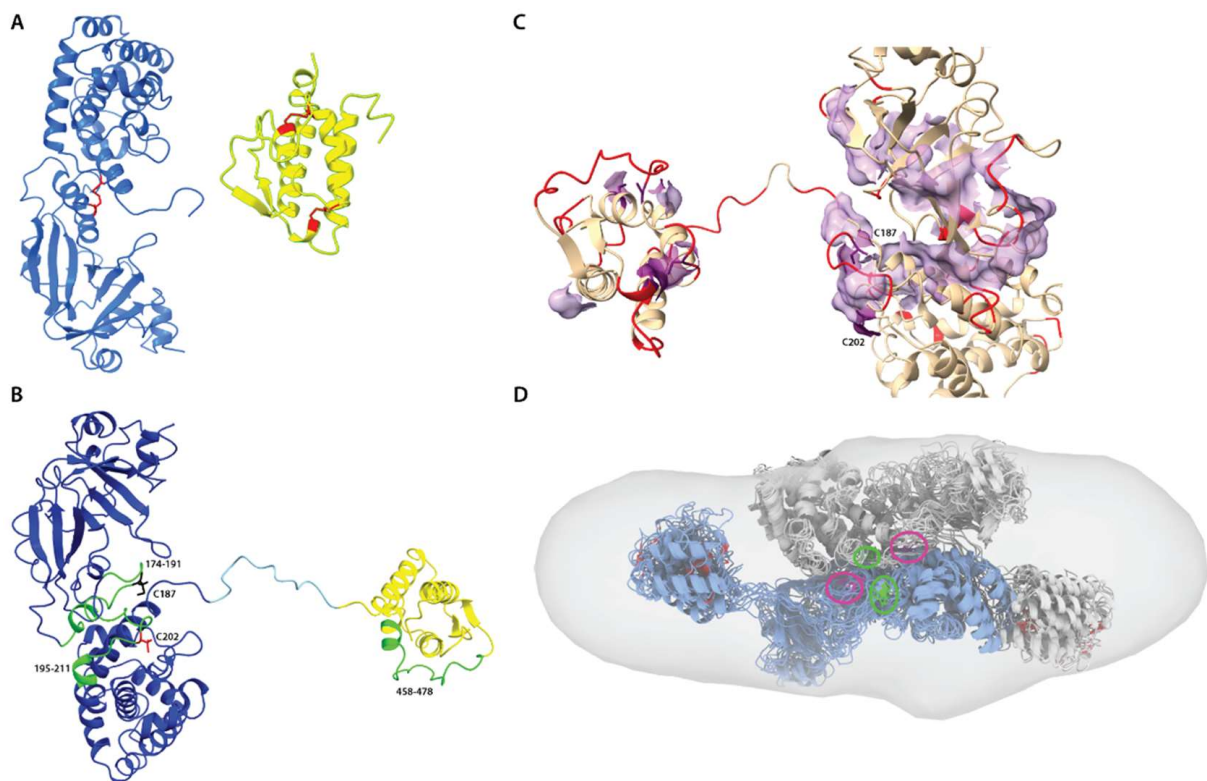
Figure 8. Disulfide mapping and protein-protein recognition. (**A**). Native disulfide bridges (red) in DT390 domains (blue) and IL-13 (yellow). (**B**). Cysteine-containing peptides identified by MS (green). Reduced C187 and C202 residues are colored black and red, respectively. (**C**). Fragment of DT390 with flexible residues colored red and surface of the major hydrophobic path. Residues C187 and C202 are highlighted in purple and belong to the hydrophobic path. (**D**). Fitting conformational assemblies of DT390-IL-13 dimer to ED model. Each "monomer" is a structural alignment of 10 most probable conformers, and both molecules are manually docked to achieve proximity between C187 and C202 residues highlighted in green and purple, respectively.

The best-scored cystine crosslink for all samples was the toxin-native one, C187–C202. Non-native linkages, for example, C187–C187, C202–C202, C187–C463, or C202–C463, were also detected but with the collective Andromeda score lower by approximately one order of magnitude (Figure S20, Supplementary Materials). This shows that the native C187–C202 bond is preserved in the oligomers and might potentially mediate linking the monomeric units. The IL-13-native crosslinks, C421–C449 and C437–C463, were poorly detected in the monomer sample. Due to the low abundance of this peptide, we cannot assuredly confirm whether these cysteines might be involved in intermolecular disulfide bonding only relying on MS data. Native disulfide network is essential for both protein folding, structure, and

biological function. Since the formation of disulfide bonds is reversible, it can be considered an environmentally sensitive switch regulating protein function. Some estimations show that 4–7% of all structurally defined disulfides in PDB are possibly allosteric. The exposed disulfide bond in the diphtheria region may be classified as the allosteric type, whereas two IL-13 disulfides belong to the structural type. On the other hand, reduction in the single cystine crosslink C187–C202, as was confirmed by our studies on the purified isolated DT390 protein (Figure S1, Supplementary Materials), also triggers aggregation. It is highly plausible that oligomerization mainly occurs during the renaturation step, due to stochastic conformational changes resulting in cysteine mispairing. In an innate biological environment, disulfide bond formation is also regulated by other proteins and changes in redox potential, which is often difficult to emulate in vitro.

For a better insight into the possible dimerization mechanism, we also evaluated the DT390-IL-13 propensity to aggregation and flexibility using computational methods (see Materials and Methods for details). Despite a relatively low aggregation score for the structure, a clear hydrophobic path can be spotted near the Cys residues involved in dimerization (Figure 8C and Figure S21, Supplementary Materials). Moreover, many surface residues are labile, including C202 and Cys residues belonging to the IL-13 domain. A dimer composed of conformational ensembles of DT390-IL-13 calculated independently from any SAXS-based methods and positioned towards the residues involved in cross-linking also can be well fitted into ED reconstruction (Figure 8D). Hence, it may be assumed that in the first step of the dimerization, non-covalent interactions driven by hydrophobic effect orient both monomers in a way that facilitates labile C202 residue to react with C187 residue resulting in the formation of a disulfide bond. The resulting dimers still possess a hydrophobic spot onto the DT390 region and conformationally labile IL-13, allowing further oligomerization and aggregating formation.

3.8. Bioactivity Functional Assays—ELISA and MTS

We analyzed the receptor-binding affinities and cytotoxic potential of the dimer and the oligomers to assess the influence of oligomerization. Sigmoidal and quadratic regression analysis of the binding ELISA with the extracellular domain of IL-13RA2 human receptor provided both EC_{50} coefficients and apparent K_d (Figure 9).
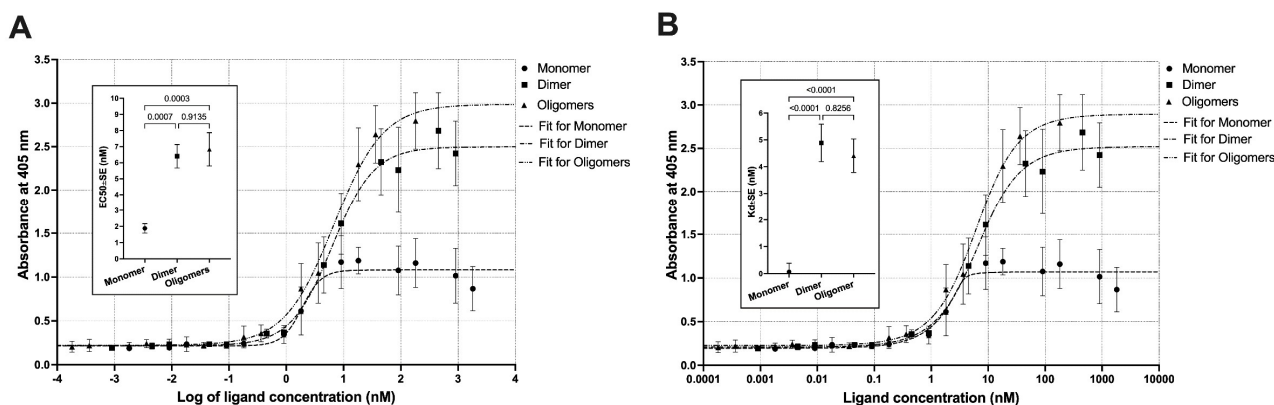
Figure 9. Non-linear regression of the ELISA experimental data for the titration of IL13RA2 receptor with the cytotoxin monomer, dimer, and oligomers. Two models were used: (**A**) sigmoidal, to derive EC_{50}, and (**B**) quadratic, to derive apparent K_d values. The inlet graphs present the fit parameters, EC_{50} or K_d, and *p*-values for the pairwise comparisons by one-way ANOVA.

Oligomerization caused an increase in EC_{50} from 1.9 ± 0.3 nM for the monomer to 6.4 ± 0.7 nM and 6.8 ± 1.0 nM for the dimer and oligomers, respectively. Our results for the monomer are in line with the IL-13RA2 affinities for IL-13 mutants reported earlier [64,65]. However, these findings also indicate that the affinity is significantly reduced by dimerization and oligomerization. The probable explanation are steric hindrances restraining

receptor binding or misfolding of the IL-13 domain in the soluble oligomers. We also tried to determine K_d derived from the quadratic model, which reflects binding thermodynamics, assuming one-to-one interaction (Equation (2)).

$$K_d = \frac{\left[R_f\right]\left[L_f\right]}{[RL]} \; f \text{ stands for } free \tag{2}$$

K_d values calculated using Equation (2) are significantly lower for the monomer (67.4 pM) compared to EC_{50} derived from the sigmoidal model, and K_d for both dimers and oligomers are similar to EC_{50}. Nevertheless, the monomer K_d has a large prediction error, making it more difficult to interpret than EC_{50} (see [66] for technical details).

The cytotoxic potential was assessed on two cell lines, U-251 with a high IL-13-RA2 externalization and LN229 with a moderate one (Figure 10 and Figure S22, Supplementary Materials). As could be expected, the monomeric fusion cytotoxin was highly effective and selective in reducing the cell viability with indeterminate IC_{50}, though it was definitely lower than 0.1 ng (1.8 pM). The cytotoxic dimer potency was lower with IC_{50}, which could be approximately 0.1 ng/mL (1.8 pM). IC_{50} of the first oligomeric fraction could be predicted to be about 1 ng (18.1 pM), so one order of magnitude lower than that of the dimer. The increase in oligomeric stoichiometries was correlated with the steady decrease in the cytotoxic potential, and fractions of the oligomeric Peak 3 containing highly scattering and unstable particles had no effect on the cell viability in the applied concentration range. It is worth noting that the cytotoxic potency for all the tested fractions was explicitly dependent on targeting IL-13RA2, since the LN229 cell line was not affected. The viability assay results are in accordance with the previous studies, with IC_{50} ranging from 0.04 to 0.23 ng/mL for the U-251 cell line [67]. Noticeably, the reduction in measured affinities is not precisely correlated with the reduction in cytotoxic potency. Both dimers and oligomers had similar, relatively low IL-13RA2 affinities. On the contrary, the evaluated cytotoxic potential decreased steadily with the increasing stoichiometries. The conclusion is that the dimers and the oligomers preserve their bioactivity to some extent. Progressive oligomerization likely leads not only to suppression of IL-13 but also DT390.
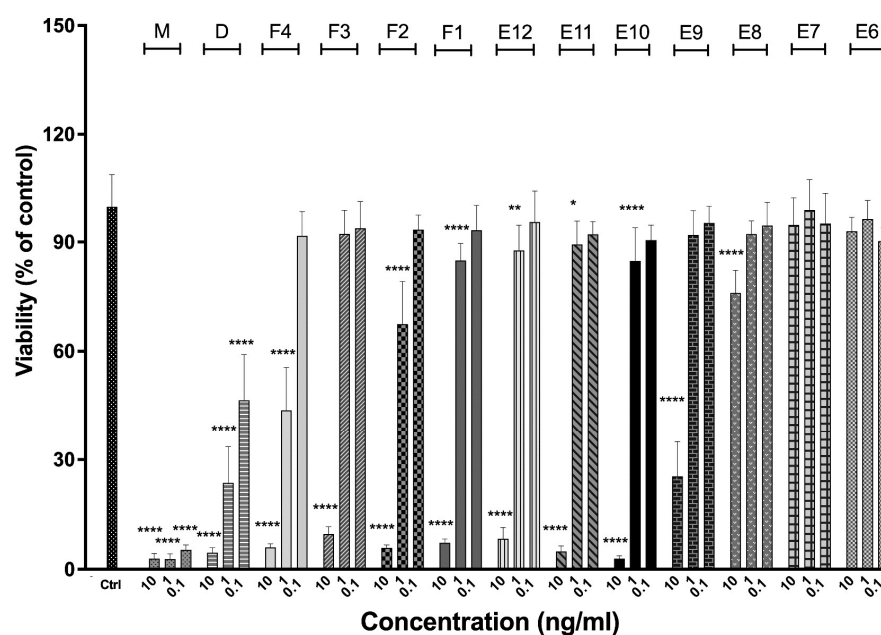
Figure 10. The viability of U-251 glioblastoma cell line in presence of the cytotoxin monomer (M), dimer (D), and oligomers (F4-E6), measured by the MTS assay, in three concentrations, 0.1, 1.0, and 10.0 ng (which correspond to 1.8, 18.1, and 181.3 pM concentration, respectively). Significant differences are indicated as: for * $p < 0.05$, ** $p < 0.01$ and **** for $p < 0.001$.

4. Conclusions

Our biochemical and structural investigation suggests that the oligomerization process of the DT390-IL-13 cytotoxin is spontaneously initiated during the refolding of the denatured protein. The resulting dimers and oligomers with lower M_w are soluble, spheroidal, and flexible particles with defined shape and size parameters. They are stabilized primarily by the native- or non-native-type disulfide linkage(s). This fact and the identified polymerization pattern for the smaller oligomers suggest some specificity of the process. Interestingly, the oligomerization mechanism involves intermolecular covalent disulfide bridging and probably weak non-covalent intermolecular surface interactions. The question remains on their comparative role and significance on both initialization and subsequent stabilization of the dimers and oligomers. Reduced receptor-affinity and cytotoxic potential indicate that dimerization and further oligomerization may impair IL-13 ligand binding with IL-13RA2 receptor or physiological functions of the toxin molecule, including efficient toxin translocation into the cell cytosol, efficient reduction of a disulfide bond in the cytoplasm, and ADP-ribosylation of translation elongation factor. However, generally, the oligomerization diminishes the efficiency of manufacturing and quality of the cytotoxic protein preparation and should be reduced. Our results indicate that blocking intermolecular interactions and promoting intramolecular ones during the refolding process would potentially shift the equilibrium from the dimeric/oligomeric toward monomeric cytotoxin. This emphasizes the need for multidimensional optimization of many parameters, including ionic strength, pH, osmolality, and amino-acid substitutions. Whether refolding conditions could be optimized to further shift the equilibrium from intermolecular toward intramolecular association is an open question. Such an endeavor would require a significant body of work since, in most situations, a trial-and-error approach is the only viable option. The optimization of the fusion proteins production process and more efficient monomer fraction isolation is currently under scrutiny in our laboratory.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/biom12081111/s1, 1. Brief description of SVD-EFA method; Figure S1: TCEP reduction of DT390 protein; Figure S2: IEX and SEC chromatography of the Exo protein; Figure S3: Results of DLS and MALS measurements for BSA and oligomeric fractions of DT390-IL-13; Figure S4: SAXS scattering profiles for all SEC-FPLC fractions; Figure S5: Exemplary SVD-EFA results; Figure S6: P(r) functions for all SVDs; Figure S7: Guinier and dimensionless Kratky plots for selected datasets; Figure S8: Flexibility analysis of selected datasets; Figure S9: Volumetric analysis for selected datasets; Figure S10: P(r) functions and the errors of BIFT for selected datasets; Figure S11: Quality assessment of ED reconstructions; Figure S12: Quality assessment of ED reconstructions; Figure S13: SAXS structural modeling of Exo-construct; Figure S14: Ensemble analysis for DT390-IL-13; Figure S15: Ensemble analysis for Exo protein; Figure S16: FoXSDock rigid docking; Figure S17: Approximate models of endo-dimer and exo-dimer; Figure S18: Estimate structure of the proposed trimer; Figure S19: Proposed structure of 10-mer based on SAXS data; Figure S20: MS analysis of DT390-IL-13; Figure S21: Computational analysis of flexibility and propensity for aggregation for DT390-IL-13; Figure S22: The viability of LN229 glioblastoma cell line in the presence of the cytotoxin monomer, dimer and oligomers; Table S1: The hydrodynamic parameters for the generated fusion cytotoxin models and the diphtheria toxin model calculated with the HullRad algorithm; Table S2: Non-linear regression parameters for the ELISA titration of the IL-13RA2 receptor with the cytotoxin monomer, dimer and oligomers; Table S3: SAXS-derived parameters for selected datasets (ESI); Table S4: Additional parameters calculated for selected datasets; Table S5: DENNS models characterization; Table S6: Characterization of conformational ensembles and oligomers for selected datasets.

Author Contributions: Conceptualization, M.M., M.Z. and B.P.; methodology, M.M., M.Z. and J.A.; software, M.M. and M.Z.; formal analysis, M.M., M.Z., M.O., R.B., A.K. and J.A.; investigation, M.M., M.Z. and B.P.; resources, M.M. and M.Z.; data curation, M.M. and M.Z.; writing—original draft preparation, M.M., M.Z. and B.P.; writing—review and editing, M.M., M.Z., M.O., R.B., A.K. and B.P.; visualization, M.M., M.Z. and A.K.; supervision, B.P.; project administration, B.P.; funding acquisition, B.P. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study, in the collection, analyses, or interpretation of data, in the writing of the manuscript, or in the decision to publish the results.

References

1. Zahaf, N.-I.; Schmidt, G. Bacterial Toxins for Cancer Therapy. *Toxins* **2017**, *9*, 236. [CrossRef] [PubMed]
2. Peigneur, S.; Tytgat, J. Toxins and drug discovery. *Toxins* **2018**, *10*, 126. [CrossRef] [PubMed]
3. Garashchenko, B.L.; Korsakova, V.A.; Yakovlev, R.Y. Radiopharmaceuticals Based on Alpha Emitters: Preparation, Properties, and Application. *Phys. At. Nucl.* **2018**, *81*, 1515–1525. [CrossRef]
4. Pagliaro, L.C.; Liu, B.; Munker, R.; Andreeff, M.; Freireich, E.J.; Scheinberg, D.A.; Rosenblum, M.G. Humanized M195 monoclonal antibody conjugated to recombinant gelonin: An anti-CD33 immunotoxin with antileukemic activity. *Clin. Cancer Res.* **1998**, *4*, 1971–1976. [PubMed]
5. Chandramohan, V.; Sampson, J.H.; Pastan, I.; Bigner, D.D. Toxin-Based Targeted Therapy for Malignant Brain Tumors. *Clin. Dev. Immunol.* **2012**, *2012*, 480429. [CrossRef] [PubMed]
6. Sedighi, M.; Zahedi Bialvaei, A.; Hamblin, M.R.; Ohadi, E.; Asadi, A.; Halajzadeh, M.; Lohrasbi, V.; Mohammadzadeh, N.; Amiriani, T.; Krutova, M.; et al. Therapeutic bacteria to combat cancer; current advances, challenges, and opportunities. *Cancer Med.* **2019**, *8*, 3167–3181. [PubMed]
7. Kaminetzky, D.; Hymes, K.B. Denileukin diftitox for the treatment of cutaneous T-cell lymphoma. *Biol. Targets Ther.* **2008**, *2*, 717–724.
8. Madhumathi, J.; Sridevi, S.; Verma, R.S. Novel TNF-related Apoptotic-inducing Ligand-based Immunotoxin for Therapeutic Targeting of CD25 Positive Leukemia. *Target. Oncol.* **2016**, *11*, 535–547. [CrossRef]
9. Frankel, A.E.; Woo, J.H.; Ahn, C.; Foss, F.M.; Duvic, M.; Neville, P.H.; Neville, D.M. Resimmune, an anti-CD3e recombinant immunotoxin, induces durable remissions in patients with cutaneous T-cell lymphoma. *Haematologica* **2015**, *100*, 794–800. [CrossRef]
10. Shan, L.; Liu, Y.; Wang, P. Recombinant Immunotoxin Therapy of Solid Tumors: Challenges and Strategies. *J. Basic Clin. Med.* **2013**, *2*, 1–6.
11. Zhu, S.; Liu, Y.; Wang, P.C.; Gu, X.; Shan, L. Recombinant immunotoxins therapy of glioblastoma: Smart design, key findings and specific challenges. *Biomed. Res. Int.* **2017**, *2017*, 7929286. [CrossRef]
12. Saito, R.; Tominaga, T. Convection-enhanced Delivery of Therapeutics for Malignant Gliomas. *Neurol. Med.-Chir.* **2017**, *57*, 8–16. [CrossRef]
13. Rossmeisl, J.H.; Herpai, D.; Quigley, M.; Cecere, T.E.; Robertson, J.L.; D'Agostino, R.B.; Hinckley, J.; Tatter, S.B.; Dickinson, P.J.; Debinski, W. Phase I trial of convection-enhanced delivery of IL13RA2 and EPHA2 receptor targeted cytotoxins in dogs with spontaneous intracranial gliomas. *Neuro-Oncol.* **2021**, *23*, 422–434. [CrossRef]
14. Zapadka, K.L.; Becher, F.J.; Gomes Dos Santos, A.L.; Jackson, S.E. Factors affecting the physical stability (aggregation) of peptide therapeutics. *Interface Focus* **2017**, *7*, 20170030. [CrossRef]
15. Vazquez, E.; Corchero, J.L.; Villaverde, A. Post-production protein stability: Trouble beyond the cell factory. *Microb. Cell Factories* **2011**, *10*, 60. [CrossRef]
16. Cromwell, M.E.M.; Hilario, E.; Jacobson, F. Protein aggregation and bioprocessing. *AAPS J.* **2006**, *8*, E572–E579. [CrossRef]
17. Wang, W.; Nema, S.; Teagarden, D. Protein aggregation—Pathways and influencing factors. *Int. J. Pharm.* **2010**, *390*, 89–99. [CrossRef]
18. Munishkina, L.A.; Ahmad, A.; Fink, A.L.; Uversky, V.N. Guiding Protein Aggregation with Macromolecular Crowding. *Biochemistry* **2008**, *47*, 8993–9006. [CrossRef]

19. Musiani, F.; Giorgetti, A. Chapter two—Protein aggregation and molecular crowding: Perspectives from multiscale simulations. *Int. Rev. Cell Mol. Biol.* **2017**, *329*, 49–77.

20. Roberts, C.J. Therapeutic protein aggregation: Mechanisms, design, and control. *Trends Biotechnol.* **2014**, *32*, 372–380. [CrossRef]

21. Lebendiker, M.; Danieli, T. Production of prone-to-aggregate proteins. *FEBS Lett.* **2014**, *588*, 236–246. [CrossRef]

22. Brange, J.; Andersen, L.; Laursen, E.D.; Meyn, G.; Rasmussen, E. Toward Understanding Insulin Fibrillation. *J. Pharm. Sci.* **1997**, *86*, 517–525. [CrossRef]

23. Rouby, G.; Tran, N.T.; Leblanc, Y.; Taverna, M.; Bihoreau, N. Investigation of monoclonal antibody dimers in a final formulated drug by separation techniques coupled to native mass spectrometry. *mAbs* **2020**, *12*, 1781743. [CrossRef]

24. Kijanka, G.; Bee, J.S.; Schenerman, M.A.; Korman, S.A.; Wu, Y.; Slutter, B.; Jiskoot, W. Monoclonal antibody dimers induced by low pH, heat or light exposure are immunogenic upon subcutaneous administration in a mouse model. *J. Pharm. Sci.* **2020**, *109*, 730–738. [CrossRef]

25. Weinfurtner, D. Analysis of disulfide bond formation in therapeutic proteins. In *Oxidative Folding of Proteins: Basic Principles, Cellular Regulation and Engineering*; Feige, M.J., Ed.; Royal Society of Chemistry: Cambridge, UK, 2018; pp. 81–98.

26. Rabdano, S.O.; Izmailov, S.A.; Luzik, D.A.; Groves, A.; Podkorytov, I.S.; Skrynnikov, N.R. Onset of disorder and protein aggregation due to oxidation-induced intermolecular disulfide bonds: Case study of RRM2 domain from TDP-43. *Sci. Rep.* **2017**, *7*, 11161. [CrossRef]

27. Rombouts, I.; Lagrin, K.A.; Scherf, K.A.; Lambrecht, M.A.; Koehler, P.; Decour, J.A. Formation and reshuffling of disulfide bonds in bovine serum albumin demonstrated using tandem mass spectrometry with collision-induced and electron-transfer dissociation. *Sci. Rep.* **2015**, *5*, 12210. [CrossRef]

28. Montoliu-Gaya, L.; Esquerda-Canals, G.; Bronsoms, S.; Villegas, S. Production of an anti-Ab antibody fragment in Pichia pastoris and in vitro and in vivo validation of its therapeutic. *PLoS ONE* **2017**, *12*, e0181480. [CrossRef]

29. Sharma, P.; Sonawane, P.; Herpai, D.; D'Agostino, R.; Rossmeisl, J.; Tatter, S.; Debinski, W. Multireceptor targeting of glioblastoma. *Neuro-Oncol. Adv.* **2020**, *2*, vdaa107. [CrossRef]

30. Krieger, E.; Nielsen, J.E.; Spronk, C.A.; Vriend, G. Fast empirical pKa prediction by Ewald summation. *J. Mol. Graph. Model.* **2006**, *25*, 481–486. [CrossRef]

31. Krieger, E.; Darden, T.; Nabuurs, S.B.; Finkelstein, A.; Vriend, G. Making optimal use of empirical energy functions: Force-field parametrization in crystal space. *Proteins* **2004**, *54*, 678–683. [CrossRef]

32. Fleming, P.J.; Fleming, K.G. HullRad: Fast Calculations of Folded and Disordered Protein and Nucleic Acid Hydrodynamic Properties. *Biophys. J.* **2018**, *114*, 856–869. [CrossRef] [PubMed]

33. Nielsen, S.S.; Toft, K.N.; Snakenborg, D.; Jeppesen, M.G.; Jacobsen, J.K.; Vestergaard, B.; Kutter, J.P.; Arleth, L. BioXTAS RAW, a software program for high-throughput automated small-angle X-ray scattering data reduction and preliminary analysis. *J. Appl. Crystallogr.* **2009**, *42*, 959–964. [CrossRef]

34. Hopkins, J.B.; Gillilan, R.E.; Skou, S. BioXTAS RAW: Improvements to a free open-source program for small-angle X-ray scattering data reduction and analysis. *J. Appl. Crystallogr.* **2017**, *50*, 1545–1553. [CrossRef] [PubMed]

35. Meisburger, S.P.; Taylor, A.B.; Khan, C.A.; Zhang, S.; Fitzpatrick, P.F.; Ando, N. Domain Movements upon Activation of Phenylalanine Hydroxylase Characterized by Crystallography and Chromatography-Coupled Small-Angle X-ray Scattering. *J. Am. Chem. Soc.* **2016**, *138*, 6506–6516. [CrossRef]

36. Manalastas-Cantos, K.; Konarev, P.V.; Hajizadeh, N.R.; Kikhney, A.G.; Petoukhov, M.V.; Molodenskiy, D.S.; Panjkovich, A.; Mertens, H.D.; Gruzinov, A.; Borges, C.; et al. ATSAS 3.0: Expanded functionality and new tools for small-angle scattering data analysis. *J. Appl. Crystallogr.* **2021**, *54*, 343–355. [CrossRef]

37. Hansen, S. Bayesian estimation of hyperparameters for indirect Fourier transformation in small-angle scattering. *J. Appl. Crystallogr.* **2000**, *33*, 1415–1421. [CrossRef]

38. Piiadov, V.; Ares de Araujo, E.; Neto, M.O.; Craievich, A.F.; Polikarpov, I. SAXSMoW 2.0: Online calculator of the molecular weight of proteins in dilute solution from experimental SAXS data measured on a relative scale. *Protein Sci.* **2019**, *28*, 454–463. [CrossRef]

39. Rambo, R.P.; Tainer, J. Accurate assessment of mass, models and resolution by small-angle scattering. *Nature* **2013**, *496*, 477–481. [CrossRef]

40. Franke, D.; Jeffries, C.M.; Svergun, D.I. Machine Learning Methods for X-ray Scattering Data Analysis from Biomacromolecular Solutions. *Biophys. J.* **2018**, *114*, 2485–2492. [CrossRef]

41. Hajizadeh, N.R.; Franke, D.; Jeffries, C.M.; Svergun, D.I. Consensus Bayesian assessment of protein molecular mass from solution X-ray scattering data. *Sci. Rep.* **2018**, *8*, 7204. [CrossRef]

42. Grant, T.D. Ab initio electron density determination directly from solution scattering data. *Nat. Methods* **2018**, *15*, 191–193. [CrossRef]

43. Pettersen, E.F.; Goddard, T.D.; Huang, C.C.; Couch, G.S.; Greenblatt, D.M.; Meng, E.C.; Ferrin, T.E. UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* **2004**, *25*, 1605–1612. [CrossRef]

44. Šali, A.; Blundell, T.L. Comparative Protein Modelling by Satisfaction of Spatial Restraints. *J. Mol. Biol.* **1993**, *234*, 779–781. [CrossRef]

45. Shapovalov, M.V.; Dunbrack, R.L., Jr. A smoothed backbone-dependent rotamer library for proteins derived from adaptive kernel density estimates and regressions. *Structure* **2011**, *19*, 844–858. [CrossRef]

46. Feig, M. Local Protein Structure Refinement via Molecular Dynamics Simulations with locPREFMD. *J. Chem. Inf. Model.* **2016**, *56*, 1304–1312. [CrossRef]
47. Goddard, T.D.; Huang, C.C.; Meng, E.C.; Pettersen, E.F.; Couch, G.S.; Morris, J.H.; Ferrin, T.E. UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci.* **2018**, *27*, 14–25. [CrossRef]
48. Schneidman-Duhovny, D.; Hammel, M.; Tainer, J.; Sali, A. FoXS, FoXSDock and MultiFoXS: Single-state and multi-state structural modeling of proteins and their complexes based on SAXS profiles. *Nucleic Acids Res.* **2016**, *44*, W424–W429. [CrossRef]
49. Schneidman-Duhovny, D.; Hammel, M.; Sali, A. FoXS: A web server for rapid computation and fitting of SAXS profiles. *Nucleic Acids Res.* **2010**, *38*, W540–W544. [CrossRef]
50. Cox, J.; Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **2008**, *26*, 1367–1372. [CrossRef]
51. Cox, J.; Neuhauser, N.; Michalski, A.; Scheltema, R.A.; Olsen, J.V.; Mann, M. Andromeda: A Peptide Search Engine Integrated into the MaxQuant Environment. *J. Proteome Res.* **2011**, *10*, 1794–1805. [CrossRef]
52. Kuriata, A.; Iglesias, V.; Pujols, J.; Kurcinski, M.; Kmiecik, S.; Ventura, S. Aggrescan3D (A3D) 2.0: Prediction and engineering of protein solubility. *Nucleic Acids Res.* **2019**, *47*, W300–W307. [CrossRef] [PubMed]
53. Kuriata, A.; Gierut, A.M.; Oleniecki, T.; Ciemny, M.P.; Kolinski, A.; Kurcinski, M.; Kmiecik, S. CABS-flex 2.0: A web server for fast simulations of flexibility of protein structures. *Nucleic Acids Res.* **2018**, *46*, W338–W343. [CrossRef] [PubMed]
54. Eble, J.A. Titration ELISA as a method to determine the dissociation constant of receptor ligand interaction. *J. Vis. Exp.* **2018**, *15*, 57334. [CrossRef] [PubMed]
55. Pham, N.B.; Meng, W.S. Protein aggregation and immunogenicity of biotherapeutics. *Int. J. Pharm.* **2020**, *585*, 119523. [CrossRef]
56. Owczarek, B.; Gerszberg, A.; Hnatuszko-Konka, K. A Brief Reminder of Systems of Production and Chromatography-Based Recovery of Recombinant Protein Biopharmaceuticals. *BioMed Res. Int.* **2019**, *2019*, 4216060. [CrossRef]
57. Humer, D.; Spadiut, O. Wanted: More monitoring and control during inclusion body processing. *World J. Microbiol. Biotechnol.* **2018**, *34*, 158. [CrossRef]
58. Klema, V.; Laub, A. The singular value decomposition: Its computation and some applications. *IEEE Trans. Autom. Control* **1980**, *25*, 164–176. [CrossRef]
59. Maeder, M. Evolving factor analysis for the resolution of overlapping chromatographic peaks. *Anal. Chem.* **1987**, *59*, 527–530. [CrossRef]
60. Kikhney, A.G.; Svergun, D.I. A practical guide to small angle X-ray scattering (SAXS) of flexible and intrinsically disordered proteins. *FEBS Lett.* **2015**, *589*, 2570–2577. [CrossRef]
61. Rambo, R.P.; Tainer, J.A. Characterizing flexible and intrinsically unstructured biological macromolecules by SAS using the Porod-Debye law. *Biopolymers* **2011**, *95*, 559–571. [CrossRef]
62. Pelikan, M.; Hura, G.L.; Hammel, M. Structure and flexibility within proteins as identified through small angle X-ray scattering. *Gen. Physiol. Biophys.* **2009**, *28*, 174–189. [CrossRef]
63. Panjkovich, A.; Svergun, D.I. Deciphering conformational transitions of proteins by small angle X-ray scattering and normal mode analysis. *Phys. Chem. Chem. Phys.* **2016**, *18*, 5707–5719. [CrossRef]
64. Madhankumar, A.B.; Mintz, A.; Debinski, W. Interleukin 13 mutants of enhanced avidity toward the glioma-associated receptor, IL13Ralpha2. *Neoplasia* **2004**, *6*, 15–22. [CrossRef]
65. Lupardus, P.J.; Birnbaum, M.E.; Garcia, K.C. Molecular basis for shared cytokine recognition revealed in the structure of an unusually high affinity complex between IL-13 and IL-13Ralpha2. *Structure* **2010**, *18*, 332–342. [CrossRef]
66. Jarmoskaite, I.; Alsadhan, I.; Vaidyanathan, P.P.; Herschlag, D. How to measure and evaluate binding affinities. *eLife* **2020**, *9*, e57264. [CrossRef]
67. Nash, K.T.; Thompson, J.P.; Debinski, W. Molecular targeting of malignant gliomas with novel multiply-mutated interleukin 13-based cytotoxins. *Crit. Rev. Oncol. Hematol.* **2001**, *39*, 87–98. [CrossRef]