



Communication

An In Silico Methodology That Facilitates Decision Making in the Engineering of Nanoscale Protein Materials

Eloi Parladé ^{1,2,*}, Eric Voltà-Durán ^{1,2,3} , Olivia Cano-Garrido ⁴, Julieta M. Sánchez ^{1,2,3,5,6}, Ugutz Unzueta ^{1,3,7} , Hèctor López-Laguna ^{1,2,3}, Naroa Serna ⁴, Montserrat Cano ⁴, Manuel Rodríguez-Mariscal ⁴, Esther Vazquez ^{1,2,3} and Antonio Villaverde ^{1,2,3,*}

- ¹ CIBER de Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), C/Monforte de Lemos 3-5, 28029 Madrid, Spain; eric.volta@uab.cat (E.V.-D.); jsanchezqa@gmail.com (J.M.S.); uunzueta@sanpau.cat (U.U.); hector.lopez@uab.cat (H.L.-L.); esther.vazquez@uab.es (E.V.)
 - ² Institut de Biotecnologia i de Biomedicina, Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain
 - ³ Departament de Genètica i de Microbiologia, Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain
 - ⁴ Nanoligent S.L., Eureka Building, Av. de Can Doménech s/n, Campus de la UAB, 08193 Bellaterra, Spain; olivia.cano.garrido@gmail.com (O.C.-G.); naroas@nanoligent.com (N.S.); mcano@nanoligent.com (M.C.); mrm@nanoligent.com (M.R.-M.)
 - ⁵ Departamento de Química, Cátedra de Química Biológica, Facultad de Ciencias Exactas, Físicas y Naturales, ICTA, Universidad Nacional de Córdoba, Av. Vélez Sársfield 1611, Córdoba 5016, Argentina
 - ⁶ Instituto de Investigaciones Biológicas y Tecnológicas (IIByT), CONICET-Universidad Nacional de Córdoba, Córdoba 5016, Argentina
 - ⁷ Biomedical Research Institute Sant Pau (IIB Sant Pau), Sant Antoni Ma Claret 167, 08025 Barcelona, Spain
- * Correspondence: eloi.parlade@uab.cat (E.P.); antonio.villaverde@uab.cat (A.V.)



Citation: Parladé, E.; Voltà-Durán, E.; Cano-Garrido, O.; Sánchez, J.M.; Unzueta, U.; López-Laguna, H.; Serna, N.; Cano, M.; Rodríguez-Mariscal, M.; Vazquez, E.; et al. An In Silico Methodology That Facilitates Decision Making in the Engineering of Nanoscale Protein Materials. *Int. J. Mol. Sci.* **2022**, *23*, 4958. <https://doi.org/10.3390/ijms23094958>

Academic Editors: Mihai V. Putz and Yi Cao

Received: 1 April 2022

Accepted: 28 April 2022

Published: 29 April 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Under the need for new functional and biocompatible materials for biomedical applications, protein engineering allows the design of assemblable polypeptides, which, as convenient building blocks of supramolecular complexes, can be produced in recombinant cells by simple and scalable methodologies. However, the stability of such materials is often overlooked or disregarded, becoming a potential bottleneck in the development and viability of novel products. In this context, we propose a design strategy based on in silico tools to detect instability areas in protein materials and to facilitate the decision making in the rational mutagenesis aimed to increase their stability and solubility. As a case study, we demonstrate the potential of this methodology to improve the stability of a humanized scaffold protein (a domain of the human nidogen), with the ability to oligomerize into regular nanoparticles usable to deliver payload drugs to tumor cells. Several nidogen mutants suggested by the method showed important and measurable improvements in their structural stability while retaining the functionalities and production yields of the original protein. Then, we propose the procedure developed here as a cost-effective routine tool in the design and optimization of multimeric protein materials prior to any experimental testing.

Keywords: nanomaterials; protein stability; nanomedicine; mutagenesis

1. Introduction

The acquisition and grouping of desired biological activities, which can be combined in a single polypeptide chain through multidomain recruitment, allows producing modular proteins with specific functionalities that are not occurring in nature [1–6]. Designing multidomain proteins is a strategy especially appealing in the generation of protein materials at micro- and nanoscales, in which the building blocks must gain assembling abilities for regulatable oligomerization [7,8]. Under urgent demands from biomedical fields such as drug delivery [9–11] and regenerative medicine [12,13], emerging protein materials offer the possibility to assemble as supramolecular structures of defined geometries and mechanical properties, combined with desired biological functionalities [14]. For instance, protein scaffolds in tissue engineering provide mechanically stable substrates for cell growth but

also display integrin-binding peptides [15] or release growth factors [16] to enhance either substrate colonization or/and proliferation. In drug delivery, protein nanoparticles offer a nanoscale size that is convenient for the prolonged circulation of payload drugs in blood and for optimal tissue and cell penetrability [17], plus specific cell targeting through solvent-display of homing peptides (i.e., specific binders of cell surface tumoral markers [18]). However, while functional recruitment through domain fusion can be organized on the simple basis of requested bio-physical properties [4,19], the whole functionality, stability, and productivity of the target proteins in cell factories are often neglected in the first design schemes. A poor or limited yield or conformational quality of recombinant proteins (being both common events [20,21]) might result in failing materials, which despite showing potential in the clinical field, are not robust enough to be brought to the industry.

In this context, rational site-directed protein engineering allows functional and structural refinement of protein functionalities that had already been achieved by domain recruitment through gene fusion. Such fine-tuning process should consider all functional and structural aspects that could be associated with the final product, with particular emphasis on the biofabrication step. The generation of structured, hierarchical tuning processes beyond the initial modular design would ensure the robust applicability of clinically relevant products, which upon a first straightforward design, could not reach optimal performance. Driven by the need to improve current protein scaffolds available for targeted drug delivery, we set out to develop a rational refining procedure described herein with transversal applicability for any kind of protein material. Specifically, we illustrate this process with the generation of a series of humanized protein nanoparticles for their application in selective drug delivery for oncology. Protein assembling, stability, and the precise cell targeting of the whole functional nanoparticles have been polished through site-directed mutagenesis of a modular polypeptide based on the human nidogen [22] empowered with T22, a peptidic ligand of the tumoral marker CXCR4 [23,24]. The architecture of the nanoparticles formed by this protein [25] and related fusion proteins [26–33] is supported by the coordination of divalent cations such as Zn^{2+} with histidine-rich domains [34], which overhang from the building blocks [35]. The rational selection of mutational targets and the sequential assay-and-error application of semi-rational protein design has rendered a fully tailored polypeptide that, in contrast to the parental protein, fulfills all the requirements for clinical development, including a smooth biofabrication process.

2. Materials and Methods

2.1. Rational Refining Procedure

The refining methodology consisted of the application of successive prediction algorithms to the structural data of the desired protein material. This was implemented in two stages. In the first one, the protein structure was subjected to a stability analysis to pinpoint areas contributing negatively to the overall robustness of the product, either by promoting precipitation or displaying excessive thermal mobility. Both issues were tackled in parallel by comparing precipitation and B-factor score predicted by Aggrescan3D 2.0 [36] and ResQ (via I-TASSER) [37,38] algorithms, respectively. High positive values in Aggrescan3D were the main indicator of instability, and residues harboring such scores were preselected for refinement, especially if they were consistent with a peak in B-factor.

In the second stage, the impact of each preselected mutation was estimated while taking into account all possible residue substitutions. The implementation of this step was performed by combining the output of DeepDDG [39], a neural network-based algorithm used to predict stability changes of protein point mutations, and STRUM [40], a machine learning algorithm trained on evolutionary information. In both stages, comparison of the scores between any of the used algorithms required a Z-score normalization of the data to cope with differences in scale of the outputs. High positive values of each index were favorable for choosing a given mutation. In the final decision regarding which residues would serve as a replacement, hydrophobic amino acids were not considered as suitable

for surface-exposed sites to avoid protein insolubility and aggregation. Reciprocally, the introduction of polar amino acids should be avoided in buried residue positions.

2.2. Protein Production and Characterization

Proteins were designed in-house, and the encoding genes were provided by GeneArt (Thermo Fisher, Waltham, MA, USA) subcloned into pET26b plasmids (Novagen-Merck, Darmstadt, Germany). The parental protein was T22-HSNBT1-H6, and all its derivatives were named T22-HSNBT2-H6 through T22-HSNBT8-H6. Protein-encoding plasmids were transformed into *Escherichia coli* BL21 DE3 (Novagen-Merck, Darmstadt, Germany), and the encoded protein was produced overnight (O/N) at 20 °C in Lysogeny Broth (LB) medium upon induction with 0.1 mM isopropyl- β -D-1-thiogalactopyranoside (IPTG). Cells were then harvested by centrifugation (15 min at 5000 \times g) and resuspended in wash buffer (20 mM Tris, 500 mM NaCl, 10 mM Imidazole, pH 8) in the presence of protease inhibitors (cOmplete™ EDTA-Free, Roche, Basel, Switzerland). Cells were then disrupted in an EmulsiFlex-C5 system (Avestin, Ottawa, ON, Canada) by 3 rounds at 8000 psi. The soluble fraction of the cell lysate containing the proteins was separated by centrifugation (45 min at 15,000 \times g) and then charged into a HisTrap HP column (GE Healthcare, Chicago, IL, USA) for purification by immobilized metal affinity chromatography (IMAC) in an ÄKTA pure system (GE Healthcare). Protein elution was achieved by applying a linear gradient of elution buffer (20 mM Tris, 500 mM NaCl, 500 mM Imidazole, pH 8). Purified protein fractions were then dialyzed against sodium carbonate (166 mM NaCO₃H, pH 8). Protein purity was determined by SDS-PAGE gel electrophoresis and subsequent Western blot immunodetection using an anti-His monoclonal antibody (Santa Cruz Biotechnology, Dallas, TX, USA). Protein integrity was also determined by MALDI-TOF mass spectrometry. The final protein concentration was determined by Bradford assay and Nanodrop.

2.3. Morphometric Characterization and Zeta Potential

Protein nanoparticle sizes were measured by Dynamic Light Scattering (DLS) at 633 nm prior to and after Zn-mediated oligomerization. Volume size distribution (expressed in nm) and zeta potential (expressed in mV) of protein materials were determined in a Zetasizer Pro (Malvern Instruments, Malvern, United Kingdom). All samples were measured in triplicates, and protein concentration was set at 1.5 mg/mL in carbonate buffer (166 mM NaCO₃H, pH 8, refractive index 1.33, viscosity 0.887 mPa·s).

2.4. Determination of Intrinsic Fluorescence and Unfolding Temperature

Fluorescence spectra were recorded in a Cary Eclipse spectrofluorometer (Agilent Technologies, Santa Clara, CA, USA). A quartz cell with a 10 mm path length and a thermostatted holder was used. The excitation and emission slits were set at 5 nm. The excitation wavelength was set at 295 nm, and the emission spectra were acquired within the 310–450 nm range. The protein concentration was 0.2 mg/mL in carbonate buffer (166 mM NaHCO₃, pH 8). In order to evaluate the conformation stability against heating, we obtained the fluorescence spectra at 5 °C temperature steps, and we calculated the center of spectral mass (CSM) for each spectrum. CSM is a weighted average of the fluorescence spectrum peak. It is also related to the relative exposure of the tryptophan (Trp) in the protein to the environment [41]. The maximum red-shift in the CSM of the Trp is compatible with large solvent accessibility and the protein unfolding. The CSM was calculated for each of the fluorescence emissions according to the following equation, where I_i is the fluorescence intensity measure at the wavelength λ_i .

$$\lambda = \frac{\sum \lambda_i \cdot I_i}{\sum I_i}$$

Midpoint unfolding temperature (T_m) was determined in GraphPad Prism 8.0.2 as the temperature value that corresponds to the inflexion point of the CSM vs. temperature curve.

2.5. Monomethyl Auristatin E Conjugation

T22-HSNBT1-H6, T22-HSNBT2-H6, and T22-HSNBT5-H6 through T22-HSNBT8-H6 were covalently linked to maleimide-functionalized Monomethyl auristatin E (MC-MMAE, 911 g/mol, Levena Biopharma, San Diego, CA, USA) through solvent-exposed cysteine-thiol and lysine-amine groups at a 1:10 protein:MMAE molar ratio upon incubation for 4 h at RT in a one-pot reaction [29]. The concentration of each protein was set to 1.5–2 mg. Afterward, all protein conjugates were dialyzed (4 rounds of 90 min) against sodium carbonate buffer in 12–14 kDa MWCO membranes (Spectra/Por™ 2) to remove unlinked MMAE and centrifuged at $15,000 \times g$ for 15 min to remove precipitated protein. Reaction efficacy was finally checked by MALDI-TOF mass spectrometry.

2.6. Protein Modeling and Statistical Analyses

Visualization and modeling of the three-dimensional structure of T22-HSNBT-H6 were conducted in ChimeraX software (version 1.2), and molecular lipophilicity potential was calculated using pyMLP [42,43]. Gaussian distribution of values was not assumed for any of the compared datasets. Hence, the nonparametric Mann–Whitney U test was used to compare means and extract significance values. All statistical analyses were performed using the built-in analysis tool in GraphPad Prism version 8.0.2.

3. Results and Discussion

3.1. In Silico Procedure

The presented rational refining procedure is herein demonstrated in oligomeric protein nanoparticles envisaged for nanomedical applications, namely the cell-targeted delivery of antitumor payload drugs. These nanoparticles result from the organization of the modular protein T22-HSNBT1-H6 with the assistance of Zn cations present in the media (Figure 1A, Patent WO2021130390). This protein is derived from the humanized version of the CXCR4-directed T22-GFP-H6 [25,44] that was re-designed, substituting GFP by the G2 domain of nidogen (Uniprot identifier: P14543) [25].

Implementation of the first step of the refining process yielded an aggregation/flexibility profile (Figure 1B) that was used to pre-select nine sites of interest to be mutated (V45, V121, F157, V176, I200, Y201, V215, V236, and L237). The first 25 residues of the protein were not eligible for mutation as they contain the ligand peptide T22 and a flexible spacer required for proper CXCR4 targeting. Four of the preselected residues were finally selected after comparison of the predicted aggregation propensity, flexibility (ResQ, B-factor), and $\Delta\Delta G$ (STRUM) scores, all stacked together in Figure 1C. Then, in the second step of the process, all possible amino acid substitutions with the potential to improve stability and solubility of the protein material were evaluated by combining $\Delta\Delta G$ scores of both STRUM and DeepDDG methods in a normalized data plot (Figure 1D). All four chosen mutations (V45T, V121Q, F157E, and V215T) involved residues exposed on the surface of the assembled protein material. Due to the hydrophilic nature of the new residues, the predicted lipophilicity was expected to decrease in comparison to the original material (Figure 1E,F). Single mutants were designed to encompass each of the best-scoring substitutions individually, and two additional mutants were designed to harbor three and all four of the selected mutations. The rationale behind the mutant with three out of the four possible mutations (T22-HSNBT6-H6) was to include as many mutations as possible without compromising the targeting properties given by T22. Then, the mutation Val215, which is the closest to T22, was not included in this construct. A modular view of the proteins and their incorporated mutations is presented in Figure 2.

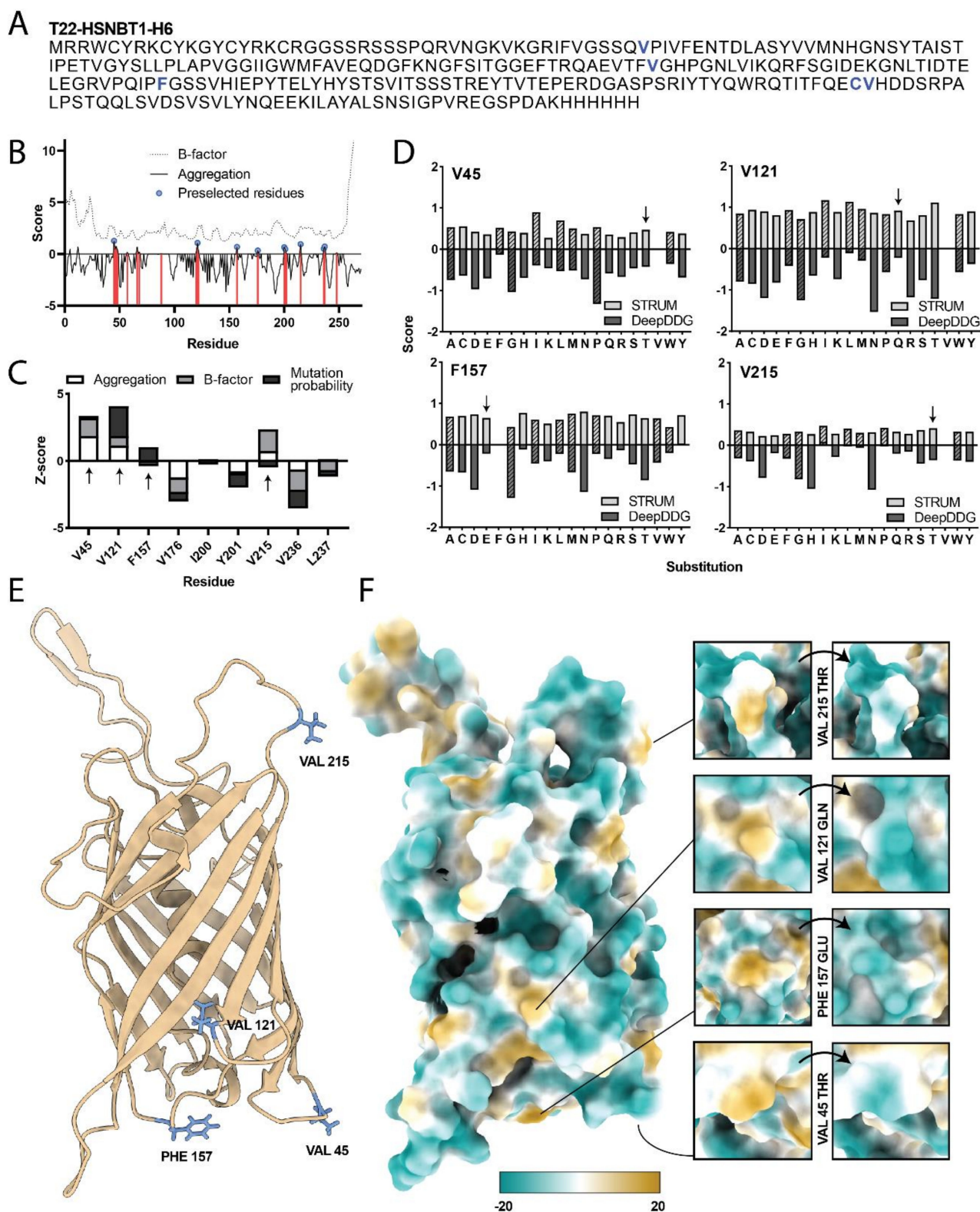


Figure 1. (A) Amino acid sequence of the construct T22–HSNBT1–H6. Residues mutated at some point are highlighted in bold and blue; numeration starts at Met0. (B) Visualization of B-factor (dotted line) and aggregation (continuous line) predictors for each residue in T22-HSNBT1-H6. Positive aggregation values indicate tendency to aggregate and are highlighted in red. Preselected residues are marked with blue dots. (C) Stacked aggregation (Aggrescan3D), B-factor (ResQ), and mutation probability (STRUM) prediction scores of the preselected residues. (D) Scores of STRUM (light gray)

and DeepDDG (dark gray) predictors for each possible substitution in the four selected residues. Hydrophobic residues are filled in a diagonal hatch pattern, and arrows indicate the chosen substitution for each amino acid. (E) Three-dimensional representation of T22-HSNBT1-H6, with the atoms of the selected residues for mutation displayed in blue. (F) Surface lipophilicity potential map. Insets on the right side highlight the predicted lipophilicity of the selected residues before and after substitution. Coloring ranges from dark cyan (most hydrophilic) to white to dark gold (most lipophilic).

T22	HSNBT1	H6	Original construct
T22	▼ HSNBT2	H6	V45T
T22	▼ HSNBT3	H6	V121Q
T22	▼ HSNBT4	H6	F157E
T22	▼ HSNBT5	H6	V215T
T22	▼ HSNBT6	H6	V45T, V121Q, F157E
T22	▼ HSNBT7	H6	V45T, V121Q, F157E, V215T
T22	▼ HSNBT8	H6	V45T, V121Q, F157E, C214S, V215T

Figure 2. Modular design of the proteins derived from T22-HSNBT1-H6. Inverted black triangles indicate the approximate location of each chosen mutation in the core module of the scaffold protein. The precise position and nature of each mutation are indicated next to the modular construct, providing the original residue, its position in the sequence, and the new residue.

An additional clone (T22-HSNBT8-H6) was also designed to incorporate all the selected mutations plus an additional substitution at C214 to remove the capacity to bind MMAE at that particular site, which was suspected of inducing protein precipitation. As this residue was already identified, only the second step of the procedure was applied to choose a substitute residue (serine). Obtaining this additional mutant was highly interesting to reduce the potential oligomeric states of the protein, envisaging highly soluble monomers still capable of forming nanoparticles upon coordination with cationic Zn [45].

In this context, we selected the I-TASSER server as the starting point of the process because, aside from integrating the ResQ algorithm, it outputs a PDB file with the predicted protein structure of the query sequence, which can be used directly as input for Aggrescan3D. Interestingly, the feasibility of using low-resolution structure models for high-accuracy stability change prediction upon point mutations has already been demonstrated [46], and a reliable tridimensional model is not required. In fact, some sequence-based methods exploiting evolutionary information show performances comparable to structure-based tools [46]. Nevertheless, new algorithms for the determination of protein stability after point mutations are constantly under development and could theoretically be used indistinctively to reproduce the current strategy. Several methods propose diverse approaches to the same end, allowing sequence or structure-based inputs [47], the incorporation of biochemical properties data [48], or the stability prediction of multiple point mutations [49]. Needless to say, protein materials with expected enzymatic or binding

capacities are susceptible to compromises in activity to favor solubility whenever a mutation tackles a catalytic residue or binding site. In such cases, users should consider the implications of proceeding with the given mutation.

3.2. Protein Expression and Purification

As stated above, the modular design of T22-HSNBT1-H6 [25] is similar to T22-GFP-H6. The GFP-based version forms multimeric nanoparticles of 12 nm efficiently targeted to CXCR4+ cancer cells [44]. Several mutated versions of the parental protein (from T22-HSNBT1-H6 to T22-HSNBT8-H6) were designed and produced in *E. coli* according to the defined purposes described above. For simplification purposes, the terms T22-HSNBT-H6 and HSNBT will be used indistinctively from now on (as synonyms) in text and figures. An initial expression test was carried out as a first approach to evaluate the feasibility of recombinant production and the quality of each new mutant protein. All mutant proteins were successfully detected by Western blot in the cell lysates, after cell disruption (Figure S1A), with molecular weights within the expected range (~30.3 kDa). Proteins containing the mutation F157E (namely HSNBT4, HSNBT6, and HSNBT7) exhibited a slightly higher migrating band. The analysis of HSNBT3 and HSNBT4 was halted at this point because an important fraction of these proteins occurred in the insoluble cell fraction (data not shown). HSNBT8, harboring the most mutations, was designed individually after verifying that its closest construct, HSNBT7, was properly produced. At this point, protein production was scaled up to 500 mL cultures maintaining the same conditions. All the proteins were successfully purified by metal affinity chromatography with suitable yields (Figure S1B), and their integrity was assessed by both SDS-PAGE/Western blot (Figure S1C) and MALDI-TOF (Figure S2A), being in 166 mM NaCO₃H buffer at pH 8. Mass spectrometry profiles confirmed that the mobility shift in SDS-PAGE of proteins harboring the mutation F157E (HSNBT6, HSNBT7, and HSNBT8) was not caused by any unexpected insertion (molar mass did not vary) and was rather due to conformational changes and/or changes in SDS binding induced by the newly accommodated glutamic acid.

3.3. Stability Study, Size Distribution, and Nanoparticle Formation

Several indicators of stability were relied upon to evaluate the properties of the new protein materials [50,51]. First, zeta potential (Figure 3A) was measured to evaluate the degree of repulsion between proteins in the dispersion and have a grasp of their stability. Generally, it is considered that high (either positive or negative) values of zeta potential are indicative of suitable stability of the system [52]. In this case, negative zeta potentials were obtained for all proteins, and while no significant differences were seen, a trend was evident in which all the mutated candidates exhibited more negative averages compared to the original protein HSNBT1. Regardless, a significant change in zeta potential was not expected as surface charges were mostly unaltered, and only one negative charge was introduced (F157E).

Second, the intrinsic fluorescence of the proteins was used to assess unfolding parameters linked to temperature [53,54]. We found that the *T_m* of the mutant proteins (Figure 3B) was in the same range as the original material yet slightly higher, indicative of increased tolerance and stability against increasing temperature. Since, among other uses, protein materials are particularly useful as drug carriers, this ability was explored by chemical conjugation with MMAE, a potent antitumor drug. Due to the hydrophobic nature of MMAE, it is essential to ensure the solubility of the conjugate, which was evaluated by assessing protein precipitation in the post-conjugation dialysis with the storage buffer (166 mM NaCO₃H, pH 8), presenting the data as a percentage soluble protein remaining relative to the original concentration (Figure 3C). Significant improvements were reported for all mutants (above 85% solubility), except for HSNBT6, which showed high variability in replicate measurements, still remaining far more soluble on average than the original material. Additionally, the cysteine removal in HSNBT8 greatly contributed to the stabilization of the protein during conjugation, as no precipitation was detected in the assays

involving this candidate. The maleimide functional group in MMAE preferably reacts with thiol groups in cysteines. Then, the only accessible cysteine not forming disulfide bonds was Cys214 (HSNBT1-HSNBT7), so conjugation at this position was deemed destabilizing. Instead, HSNBT8 had Cys214 replaced with Ser214, meaning that MMAE was distributed throughout the many amine groups from exposed lysines, keeping the protein material soluble. Supporting this claim, up to one linked MMAE was detected via MALDI-TOF in the soluble fractions of HSNBT1-HSNBT7, while up to three linked MMAE could be identified in HSNBT8, which did not precipitate (Figure S1D).

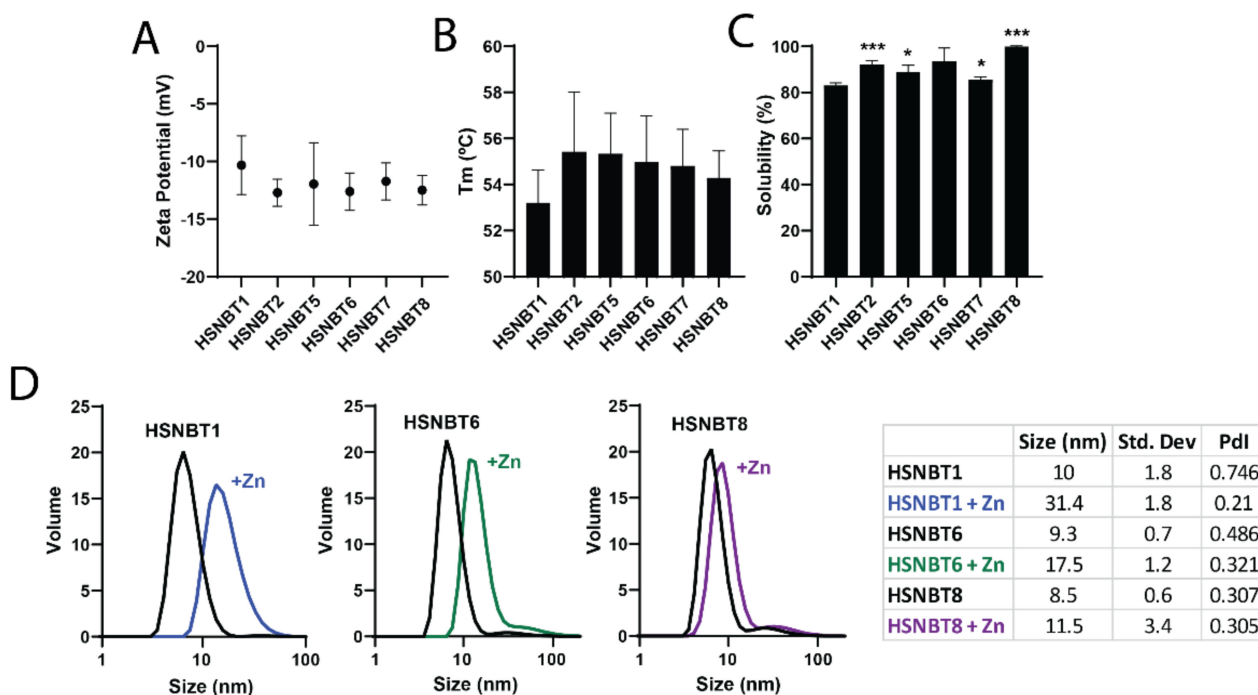


Figure 3. (A). Distribution of the zeta potential values measured for the studied proteins. (B) Tm of the studied proteins obtained from CSM curves. (C) Precipitated portion of each of the protein mutants after the conjugation with the drug MMAE. (D) Size-volume profile (left) and size data (right) of selected candidates measured by DLS before and after nanoparticle formation induced with ZnCl₂. Significance indicated as: * $p \leq 0.05$, *** $p \leq 0.001$.

At this point, HSNBT6 and HSNBT8 were selected to proceed further with the characterization because of their high production yields (namely 26 and up to 28 mg/L, respectively). Volume size distribution was studied via DLS using HSNBT1 as a reference. As depicted in Figure 3D, the same size was maintained in all protein materials, confirming that the chosen mutations did not affect the quaternary structure of the protein in its native state.

Finally, the capacity of the new protein materials to assemble into nanoparticles was explored by adding ZnCl₂ at a final concentration of 0.04–0.08 mM. In all cases, an increase in size was achieved (Figure 3D), reflecting nanoscale oligomerization. Specifically, HSNBT6 and HSNBT8 formed nanoparticles of around 17.5 and 11.5 nm, respectively, while the reference HSNBT1 yielded 31 nm-nanoparticles. Smaller nanoparticles in HSNBT8 were expected, as the lone cysteine at position 214 that was removed could promote the dimerization of the building blocks before assembly with Zn²⁺, leading to bigger nanoparticles. Nevertheless, these size ranges were within those considered optimal in medical applications, namely between 10 and 100 nm, that favor recirculation in the bloodstream and prevent renal clearance [55]. Most importantly, the results also confirm that mutations that originated from the refining protocol did not interfere with the Zn-driven assembling capacity

of the building block. Hence, both candidates HSNBT6 and HSNBT8 represent successful improvements over the starting material, suitable for clinically oriented applications.

4. Conclusions

In this study, we have presented a sequential *in silico* screening procedure to be routinely applied to improve the stability of nanoscale protein materials envisaged for biomedical applications. This platform provides an array of favorable mutations prone to stabilize the final product, either in the monomeric form or as nano-sized oligomers. While of a non-automated nature, the method offers flexibility and possibilities for further modulation since the users may follow the basic steps using their preferred algorithms, fitting specific needs or improving the whole process with updated algorithms. Thus, the inclusion of brief *in vitro* optimization methods, such as the one described herein, into protein design processes should improve the success chance of the resulting materials at virtually no cost. Moreover, adapting such methodologies will enhance project viability prior to reaching the experimental phase, where having unsuccessful candidates translates into large expenditures of time, personnel, and laboratory consumables.

5. Patents

Protein scaffolds described herein are protected under patent WO2021130390A1, licensed to Nanoligent S.L.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms23094958/s1>.

Author Contributions: Conceptualization, E.P., O.C.-G. and A.V.; methodology, E.P., E.V.-D., J.M.S. and H.L.-L.; software, E.P.; formal analysis, E.P.; investigation, E.P., O.C.-G. and N.S.; resources, U.U., E.V. and A.V.; writing—original draft preparation, E.P. and A.V.; writing—review and editing, E.P., E.V.-D., O.C.-G., J.M.S., U.U., H.L.-L., N.S., M.C., M.R.-M., E.V. and A.V.; project administration, M.C., M.R.-M., E.V. and A.V.; funding acquisition, M.C., M.R.-M., E.V. and A.V. All authors have read and agreed to the published version of the manuscript.

Funding: Nanoligent S.L. participates in the project «NANOTRAC. Desarrollo de un nuevo medicamento selectivo con potente actividad antimetastática» funded in the call Retos Colaboración of the Programa Estatal de Investigación, Desarrollo e Innovación Orientada a los Retos de la Sociedad, under the Plan Estatal de Investigación Científica y Técnica y de Innovación 2017–2020 (RTC-2017-6125-1). The authors are also indebted to the CERCA programme from la Generalitat de Catalunya, Spain, and to the Networking Research Center on Bioengineering, Biomaterials and Nanomedicine (CIBER-BBN), Spain, which is an initiative funded by the VI National R&D&I Plan 2008–2011, Iniciativa Ingenio 2010, Consolider Program, CIBER Actions and financed by the Instituto de Salud Carlos III, with assistance from the European Regional Development Fund. H.L.L. was supported by a predoctoral fellowship from AGAUR (2019FI_B00352), and E.V.D. was supported by a predoctoral fellowship from Ministerio de Ciencia e Innovación, Spain (FPU18/04615). U.U. was supported by a Miguel Servet contract (CP19/00028) from ISCIII co-funded by European Social Fund (ESF investing in your future). A.V. received an ICREA ACADEMIA award.

Data Availability Statement: The data presented in this study are openly available at: <https://www.dropbox.com/s/p5h0ogt5cf5t0rv/DATA.xlsx?dl=0> (accessed on 1 April 2022).

Acknowledgments: Molecular graphics and analyses were performed with UCSF ChimeraX, developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from the National Institutes of Health R01-GM129325 and the Office of Cyber Infrastructure and Computational Biology, National Institute of Allergy and Infectious Diseases.

Conflicts of Interest: M.R.-M., E.V. and A.V. are co-founders of Nanoligent S.L.; O.C.-G., N.S., M.C. and M.R.-M. are or were employed by Nanoligent S.L. during this study. The other authors declare no conflict of interest.

References

1. Dimarco, R.L.; Heilshorn, S.C. Multifunctional Materials through Modular Protein Engineering. *Adv. Mater.* **2012**, *24*, 3923–3940. [[CrossRef](#)] [[PubMed](#)]
2. Siciliano, V.; Diandreth, B.; Monel, B.; Beal, J.; Huh, J.; Clayton, K.L.; Wroblewska, L.; McKeon, A.; Walker, B.D.; Weiss, R. Engineering Modular Intracellular Protein Sensor-Actuator Devices. *Nat. Commun.* **2018**, *9*. [[CrossRef](#)] [[PubMed](#)]
3. Aris, A.; Villaverde, A. Modular Protein Engineering for Non-Viral Gene Therapy. *Trends Biotechnol.* **2004**, *22*, 371–377. [[CrossRef](#)] [[PubMed](#)]
4. Vazquez, E.; Ferrer-Miralles, N.; Mangues, R.; Corchero, J.; Schwartz, S., Jr.; Villaverde, A. Modular Protein Engineering in Emerging Cancer Therapies. *Curr. Pharm. Des.* **2009**, *15*, 893–916. [[CrossRef](#)]
5. Glasgow, A.A.; Huang, Y.M.; Mandell, D.J.; Thompson, M.; Ritterson, R.; Loshbaugh, A.L.; Pellegrino, J.; Krivacic, C.; Pache, R.A.; Barlow, K.A.; et al. Computational Design of a Modular Protein Sense-Response System. *Science* **2019**, *366*, 1024–1028. [[CrossRef](#)]
6. Matico, R.; Szewczuk, L.M.; Pietrak, B.; Chen, S.; Dul, E.; Bonnette, W.G.; Meinhold, D.W.; Quinque, G.; Totoritis, R.; Lewis, T.; et al. Modular Protein Ligation: A New Paradigm as a Reagent Platform for Pre-Clinical Drug Discovery. *Sci. Rep.* **2019**, *9*, 13078. [[CrossRef](#)]
7. Main, E.R.G.; Phillips, J.J.; Millership, C. Repeat Protein Engineering: Creating Functional Nanostructures/Biomaterials from Modular Building Blocks. *Biochem. Soc. Trans.* **2013**, *41*, 1152–1158. [[CrossRef](#)]
8. Vazquez, E.; Mangues, R.; Villaverde, A. Functional Recruitment for Drug Delivery through Protein-Based Nanotechnologies. *Nanomedicine* **2016**, *11*, 1333–1336. [[CrossRef](#)]
9. Korpi, A.; Anaya-Plaza, E.; Välimäki, S.; Kostianen, M. Highly Ordered Protein Cage Assemblies: A Toolkit for New Materials. *Wiley Interdiscip. Rev. Nanomed. Nanobiotechnol.* **2020**, *12*, e1578. [[CrossRef](#)]
10. Jao, D.; Xue, Y.; Medina, J.; Hu, X. Protein-Based Drug-Delivery Materials. *Materials* **2017**, *10*, 517. [[CrossRef](#)]
11. Berckman, E.A.; Hartzell, E.J.; Mitkas, A.A.; Sun, Q.; Chen, W. Biological Assembly of Modular Protein Building Blocks as Sensing, Delivery, and Therapeutic Agents. *Annu. Rev. Chem. Biomol. Eng.* **2020**, *11*, 35–62. [[CrossRef](#)] [[PubMed](#)]
12. Hussey, G.S.; Dziki, J.L.; Badylak, S.F. Extracellular Matrix-Based Materials for Regenerative Medicine. *Nat. Rev. Mater.* **2018**, *3*, 159–173. [[CrossRef](#)]
13. DeFrates, K.G.; Moore, R.; Borgesi, J.; Lin, G.; Mulderig, T.; Beachley, V.; Hu, X. Protein-Based Fiber Materials in Medicine: A Review. *Nanomaterials* **2018**, *8*, 457. [[CrossRef](#)]
14. Hamley, I.W. Protein Assemblies: Nature-Inspired and Designed Nanostructures. *Biomacromolecules* **2019**, *20*, 1829–1848. [[CrossRef](#)] [[PubMed](#)]
15. Dhavalikar, P.; Robinson, A.; Lan, Z.; Jenkins, D.; Chwatko, M.; Salhadar, K.; Jose, A.; Kar, R.; Shoga, E.; Kannapiran, A.; et al. Review of Integrin-Targeting Biomaterials in Tissue Engineering. *Adv. Healthc. Mater.* **2020**, *9*, 2000795. [[CrossRef](#)] [[PubMed](#)]
16. Subbiah, R.; Guldberg, R.E. Materials Science and Design Principles of Growth Factor Delivery Systems in Tissue Engineering and Regenerative Medicine. *Adv. Healthc. Mater.* **2019**, *8*, 1801000. [[CrossRef](#)]
17. Jain, A.; Singh, S.K.; Arya, S.K.; Kundu, S.C.; Kapoor, S. Protein Nanoparticles: Promising Platforms for Drug Delivery Applications. *ACS Biomater. Sci. Eng.* **2018**, *4*, 3939–3961. [[CrossRef](#)]
18. Hamley, I.W. Small Bioactive Peptides for Biomaterials Design and Therapeutics. *Chem. Rev.* **2017**, *117*, 14015–14041. [[CrossRef](#)]
19. Serna, N.; Sánchez-García, L.; Unzueta, U.; Díaz, R.; Vázquez, E.; Mangues, R.; Villaverde, A. Protein-Based Therapeutic Killing for Cancer Therapies. *Trends Biotechnol.* **2018**, *36*, 318–335. [[CrossRef](#)]
20. Rosano, G.L.; Ceccarelli, E.A. Recombinant Protein Expression in Escherichia Coli: Advances and Challenges. *Front. Microbiol.* **2014**, *5*, 172. [[CrossRef](#)]
21. Corchero, J.L.; Gasser, B.; Resina, D.; Smith, W.; Parrilli, E.; Vázquez, F.; Abasolo, I.; Giuliani, M.; Jäntti, J.; Ferrer, P.; et al. Unconventional Microbial Systems for the Cost-Efficient Production of High-Quality Protein Therapeutics. *Biotechnol. Adv.* **2013**, *31*, 140–153. [[CrossRef](#)] [[PubMed](#)]
22. Nagayoshi, T.; Sanborn, D.; Hickok, N.J.; Olsen, D.R.; Fazio, M.J.; Chu, M.L.; Knowlton, R.; Mann, K.; Deutzmann, R.; Timpl, R.; et al. Human Nidogen: Complete Amino Acid Sequence and Structural Domains Deduced from cDNAs, and Evidence for Polymorphism of the Gene. *DNA* **1989**, *8*, 581–594. [[CrossRef](#)] [[PubMed](#)]
23. Tamamura, H.; Kuroda, M.; Masuda, M.; Otaka, A.; Funakoshi, S.; Nakashima, H.; Yamamoto, N.; Waki, M.; Matsumoto, A.; Lancelin, J.M.; et al. A Comparative Study of the Solution Structures of Tachyplesin I and a Novel Anti-HIV Synthetic Peptide, T22 ([Tyr5,12, Lys7]-Polyphemusin II), Determined by Nuclear Magnetic Resonance. *Biochim. Biophys. Acta (BBA) Protein Struct. Mol.* **1993**, *1163*, 209–216. [[CrossRef](#)]
24. Tamamura, H.; Imai, M.; Ishihara, T.; Masuda, M.; Funakoshi, H.; Oyake, H.; Murakami, T.; Arakaki, R.; Nakashima, H.; Otaka, A.; et al. Pharmacophore Identification of a Chemokine Receptor (CXCR4) Antagonist, T22 ([Tyr5,12, Lys7]-Polyphemusin II), Which Specifically Blocks T Cell-Line-Tropic HIV-1 Infection. *Bioorganic Med. Chem.* **1998**, *6*, 1033–1041. [[CrossRef](#)]
25. Álamo, P.; Cedano, J.; Conchillo-Sole, O.; Cano-Garrido, O.; Alba-Castellon, L.; Serna, N.; Aviñó, A.; Carrasco-Diaz, L.M.; Sánchez-Chardi, A.; Martínez-Torró, C.; et al. Rational Engineering of a Human GFP-like Protein Scaffold for Humanized Targeted Nanomedicines. *Acta Biomater.* **2021**, *130*, 211–222. [[CrossRef](#)] [[PubMed](#)]
26. López-Laguna, H.; Sánchez-García, L.; Serna, N.; Voltà-Durán, E.; Sánchez, J.M.; Sánchez-Chardi, A.; Unzueta, U.; Łoś, M.; Villaverde, A.; Vázquez, E. Engineering Protein Nanoparticles Out from Components of the Human Microbiome. *Small* **2020**, *16*, 2001885. [[CrossRef](#)]

27. Serna, N.; Álamo, P.; Ramesh, P.; Vinokurova, D.; Sánchez-García, L.; Unzueta, U.; Gallardo, A.; Céspedes, M.V.; Vázquez, E.; Villaverde, A.; et al. Nanostructured Toxins for the Selective Destruction of Drug-Resistant Human CXCR4+ Colorectal Cancer Stem Cells. *J. Control. Release* **2020**, *320*, 96–104. [[CrossRef](#)]
28. López-Laguna, H.; Cubarsi, R.; Unzueta, U.; Mangues, R.; Vázquez, E.; Villaverde, A. Endosomal Escape of Protein Nanoparticles Engineered through Humanized Histidine-Rich Peptides. *Sci. China Mater.* **2020**, *63*, 644–653. [[CrossRef](#)]
29. Pallarès, V.; Unzueta, U.; Falgàs, A.; Sánchez-García, L.; Serna, N.; Gallardo, A.; Morris, G.A.; Alba-Castellón, L.; Álamo, P.; Sierra, J.; et al. An Auristatin Nanoconjugate Targeting CXCR4+ Leukemic Cells Blocks Acute Myeloid Leukemia Dissemination. *J. Hematol. Oncol.* **2020**, *13*. [[CrossRef](#)]
30. Unzueta, U.; Roldán, M.; Pesarrodonna, M.; Benitez, R.; Sánchez-Chardi, A.; Conchillo-Solé, O.; Mangues, R.; Villaverde, A.; Vázquez, E. Self-Assembling as Regular Nanoparticles Dramatically Minimizes Photobleaching of Tumour-Targeted GFP. *Acta Biomater.* **2020**, *103*, 272–280. [[CrossRef](#)]
31. Falgàs, A.; Pallarès, V.; Unzueta, U.; Céspedes, M.V.; Arroyo-Solera, I.; Moreno, M.J.; Gallardo, A.; Mangues, M.A.; Sierra, J.; Villaverde, A.; et al. A CXCR4-Targeted Nanocarrier Achieves Highly Selective Tumor Uptake in Diffuse Large B-Cell Lymphoma Mouse Models. *Haematologica* **2020**, *105*, 741–753. [[CrossRef](#)]
32. Sánchez-García, L.; Sala, R.; Serna, N.; Álamo, P.; Parladé, E.; Alba-Castellón, L.; Voltà-Durán, E.; Sánchez-Chardi, A.; Unzueta, U.; Vázquez, E.; et al. A Refined Cocktail of Pro-Apoptotic Nanoparticles Boosts Anti-Tumor Activity. *Acta Biomater.* **2020**, *113*, 584–596. [[CrossRef](#)] [[PubMed](#)]
33. Serna, N.; Cano-Garrido, O.; Sánchez-García, L.; Pesarrodonna, M.; Unzueta, U.; Sánchez-Chardi, A.; Mangues, R.; Vázquez, E.; Villaverde, A. Engineering Protein Venoms as Self-Assembling CXCR4-Targeted Cytotoxic Nanoparticles. *Part. Part. Syst. Charact.* **2020**, *37*, 2000040. [[CrossRef](#)]
34. López-Laguna, H.; Sánchez, J.; Unzueta, U.; Mangues, R.; Vázquez, E.; Villaverde, A. Divalent Cations: A Molecular Glue for Protein Materials. *Trends Biochem. Sci.* **2020**, *45*, 992–1003. [[CrossRef](#)]
35. López-Laguna, H.; Unzueta, U.; Conchillo-Solé, O.; Sánchez-Chardi, A.; Pesarrodonna, M.; Cano-Garrido, O.; Voltà, E.; Sánchez-García, L.; Serna, N.; Saccardo, P.; et al. Assembly of Histidine-Rich Protein Materials Controlled through Divalent Cations. *Acta Biomater.* **2019**, *83*, 257–264. [[CrossRef](#)] [[PubMed](#)]
36. 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](#)]
37. Yang, J.; Wang, Y.; Zhang, Y. ResQ: An Approach to Unified Estimation of B-Factor and Residue-Specific Error in Protein Structure Prediction. *J. Mol. Biol.* **2016**, *428*, 693–701. [[CrossRef](#)]
38. Roy, A.; Kucukural, A.; Zhang, Y. I-TASSER: A Unified Platform for Automated Protein Structure and Function Prediction. *Nat. Protoc.* **2010**, *5*, 725–738. [[CrossRef](#)]
39. Cao, H.; Wang, J.; He, L.; Qi, Y.; Zhang, J.Z. DeepDDG: Predicting the Stability Change of Protein Point Mutations Using Neural Networks. *J. Chem. Inf. Modeling* **2019**, *59*, 1508–1514. [[CrossRef](#)]
40. Quan, L.; Lv, Q.; Zhang, Y. STRUM: Structure-Based Prediction of Protein Stability Changes upon Single-Point Mutation. *Bioinformatics* **2016**, *32*, 2936–2946. [[CrossRef](#)]
41. Ahmad, P.; ul Islam, B.; Allarakha, S.; Dixit, K.; Ali, A.; Rabbani, G.; Khan, R.H.; Siddiqui, S.A. Preferential Recognition of Peroxynitrite-Modified Human Serum Albumin by Circulating Autoantibodies in Cancer. *Int. J. Biol. Macromol.* **2015**, *72*, 875–882. [[CrossRef](#)] [[PubMed](#)]
42. Broto, P.; Moreau, G.; Vandycke, C. Molecular Structures: Perception, Autocorrelation Descriptor and Sar Studies: System of Atomic Contributions for the Calculation of the n-Octanol/Water Partition Coefficients. *Eur. J. Med. Chem.* **1984**, *19*, 71–78.
43. Laguerre, M.; Saux, M.; Dubost, J.P.; Carpy, A. MLPP: A Program for the Calculation of Molecular Lipophilicity Potential in Proteins. *Pharm. Pharmacol. Commun.* **1997**, *3*, 217–222.
44. Céspedes, M.V.; Unzueta, U.; Aviñó, A.; Gallardo, A.; Álamo, P.; Sala, R.; Sánchez-Chardi, A.; Casanova, I.; Mangues, M.A.; Lopez-Pousa, A.; et al. Selective Depletion of Metastatic Stem Cells as Therapy for Human Colorectal Cancer. *EMBO Mol. Med.* **2018**, *10*, e8772. [[CrossRef](#)] [[PubMed](#)]
45. López-Laguna, H.; Sánchez, J.M.; Carratalá, J.V.; Rojas-Peña, M.; Sánchez-García, L.; Parladé, E.; Sánchez-Chardi, A.; Voltà-Durán, E.; Serna, N.; Cano-Garrido, O.; et al. Biofabrication of Functional Protein Nanoparticles through Simple His-Tag Engineering. *ACS Sustain. Chem. Eng.* **2021**, *9*, 12341–12354. [[CrossRef](#)] [[PubMed](#)]
46. Fariselli, P.; Martelli, P.L.; Savojardo, C.; Casadio, R. INPS: Predicting the Impact of Non-Synonymous Variations on Protein Stability from Sequence. *Bioinformatics* **2015**, *31*, 2816–2821. [[CrossRef](#)]
47. Li, B.; Yang, Y.T.; Capra, J.A.; Gerstein, M.B. Predicting Changes in Protein Thermodynamic Stability upon Point Mutation with Deep 3D Convolutional Neural Networks. *PLoS Comput. Biol.* **2020**, *16*, e1008291. [[CrossRef](#)]
48. Chen, C.W.; Lin, M.H.; Liao, C.C.; Chang, H.P.; Chu, Y.W. IStable 2.0: Predicting Protein Thermal Stability Changes by Integrating Various Characteristic Modules. *Comput. Struct. Biotechnol. J.* **2020**, *18*, 622–630. [[CrossRef](#)]
49. Montanucci, L.; Capriotti, E.; Frank, Y.; Ben-Tal, N.; Fariselli, P. DDGun: An Untrained Method for the Prediction of Protein Stability Changes upon Single and Multiple Point Variations. *BMC Bioinform.* **2019**, *20*, 335. [[CrossRef](#)]
50. Rabbani, G.; Ahmad, E.; Zaidi, N.; Fatima, S.; Khan, R.H. PH-Induced Molten Globule State of *Rhizopus Niveus* Lipase Is More Resistant Against Thermal and Chemical Denaturation Than Its Native State. *Cell Biochem. Biophys.* **2012**, *62*, 487–499. [[CrossRef](#)]

51. Rabbani, G.; Kaur, J.; Ahmad, E.; Khan, R.H.; Jain, S.K. Structural Characteristics of Thermostable Immunogenic Outer Membrane Protein from *Salmonella Enterica* Serovar Typhi. *Appl. Microbiol. Biotechnol.* **2014**, *98*, 2533–2543. [[CrossRef](#)] [[PubMed](#)]
52. Lu, G.W.; Gao, P. Emulsions and Microemulsions for Topical and Transdermal Drug Delivery. In *Handbook of Non-Invasive Drug Delivery Systems*; Elsevier: Amsterdam, The Netherlands, 2010; pp. 59–94.
53. Rabbani, G.; Ahmad, E.; Zaidi, N.; Khan, R.H. PH-Dependent Conformational Transitions in Conalbumin (Ovotransferrin), a Metalloproteinase from Hen Egg White. *Cell Biochem. Biophys.* **2011**, *61*, 551–560. [[CrossRef](#)] [[PubMed](#)]
54. Rabbani, G.; Ahmad, E.; Khan, M.V.; Ashraf, M.T.; Bhat, R.; Khan, R.H. Impact of Structural Stability of Cold Adapted *Candida Antarctica* Lipase B (CaLB): In Relation to PH, Chemical and Thermal Denaturation. *RSC Adv.* **2015**, *5*, 20115–20131. [[CrossRef](#)]
55. Longmire, M.; Choyke, P.L.; Kobayashi, H. Clearance Properties of Nano-Sized Particles and Molecules as Imaging Agents: Considerations and Caveats. *Nanomedicine* **2008**, *3*, 703–717. [[CrossRef](#)] [[PubMed](#)]