Epitope prediction and homology modeling.

T22-HSNBT-H6 monomer model was generated through Modeller 9v13 [1] as described elsewhere [2], but using 1GL4 (residues 400 to 631) as HSNBT template instead of using 1QYO. Then, prediction of putative immunogenic regions was carried out in similar way to previously published works [3-6]. Briefly: For each molecule to be predicted, 5 different structures where selected. For each structure a prediction for structural epitopes using MLCE / BEPPE [7, 8] method was performed and a consensus was extracted (all regions present in all predictions). The same procedure was followed using EDP prediction method [9] and a consensus of both methods was extracted. To select the five initial structures to analyze, from the 500 previously generated models for the T22-HSNBT-H6 monomer, the best (according to DOPE score [10]) the worse and three intermediate ones were picked. For all of them, the region corresponding to the HSNBT was extracted and FoldX [11] RepairPDB and BuildModel functions were used to obtain the final structures to be used as input for the prediction methods, for both, the Human NidogenG2 sequence and HSNBT.

T22-HSNBT-H6 challenge in an immunocompetent model

Immunogenicity derived from T22-HSNBT-H6 treatment was analyzed measuring the presence of IgG in serum by Western blot and quantifying the percentage of CD3+ cells and CD19+ cells in spleen by flow cytometry. For that, Balb/c female mice (eight weeks old) were obtained from Charles River Laboratories and housed with sterile food and water ad libitum. After one week in quarantine, Balb/c mice were divided randomly into two experimental groups. One group (Buffer; n=3) was intravenously (IV) injected with 100 μL buffer (NaCO3H 166mM + NaCl 333mM, pH 8). The second group (T22-HSNBT-H6, n=3) was administered with 100 μg of T22-HSNBT-H6. Both groups were treated with a daily dose for two days and euthanized the day after to obtain blood and spleen samples. For blood collection, mice were anesthetized prior intracardiac puncture. The serum was obtained by centrifugation and stored at -80ºC until analysis.

IgG detection by Western Blot analysis:

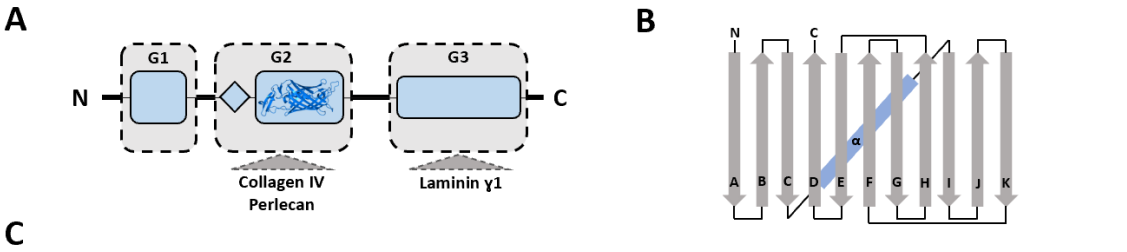
Serum samples were diluted 1:50 in water, and boiled for 3 minutes at 95ºC before loading into 10% acrylamide SDS-PAGE gels. For the Western blot, an anti-mouse IgG-HRP antibody (JACKSON IMMUNO RESEARCH EUROPE) was incubated with the nitrocellulose membrane at 1/10000 dilution at room temperature for 1 hour. Images were acquired using ChemiDoc XRS+ with Image Lab software. Bands quantification was performed using Image J software.

Spleen Flow cytometry analysis:

The spleens collected in the harvesting were processed to a single-cell preparation following previously described protocol. [12]. Single-cell suspensions were stored in liquid nitrogen until analysis. Splenocytes were stained with surface biomarkers conjugated to fluorochromes (Leukocyte marker, CD45-VioGreen; T-cell marker CD3-PerCP-Vio700; and B-cell marker CD19-PE-Vio700; Miltenyi Biotech, Cambridge, MA, USA). A viability marker, Viability Fixable 488/520 Dye (Miltenyi Biotech, Cambridge, MA, USA), was used to discriminate dead and viable cells. Single-cell population was gated based on forward scatter area, height and viability marker emission. This population was then used for immunophenotyping analysis with specific fluorochromes. The CD3+ cells percentage comes from the viable cells gate. On the other hand, CD3- CD19+ cells percentage comes from CD3- cells gate. All data collection was performed using MACSQuant Analyzer 10 (Miltenyi Biotech, Cambridge, MA, USA) and data analyses were performed using MACSQuantify software (Miltenyi Biotech, Cambridge, MA, USA). Compensation was performed using single colour controls.

Ligand binding assay.

Surface plasmon resonance assay was performed at 25 °C using a Biacore™ T200 (Cytiva) at the Scientific and Technological Centres of the University of Barcelona (CCiTUB). The ligands, collagen IV (C6745 Sigma) and perlecan (H4777 Sigma) were immobilized at different channels on a sensor chip CM5 Series S (Cytiva). Immobilization was achieved by amine coupling as described by the manufacturer. The first channel was used as reference and it was treated as the same way but without ligand. The running buffer was Phosphate Buffered Saline (PBS) pH 7.4 + 0.15M NaCl + 0.05% Tween 20. The analytes were applied in serial dilutions from 10 µM to 0.07 µM at 25ºC using flow rate of 30 µl/min. Contact time was set to 60s and dissociation curve monitored for 10 min. Association and dissociation curves were then fitted according to a 1:1 model with Biacore T200 Evaluation software (version 3.1). All experiments were performed in duplicate.



C

```
sp|P14543|NID1_HUMAN     HQQHPQVIDVDEVEETGVVFSYNTDSRQTCANNRHQCSVHAECRDYATGFCCSCVAGYTG  420
sp|P10493|NID1_MOUSE     PQHHPQVIDVDEVEETGVVFSYNTGSQQTCANNRHQCSVHAECRDYATGFCCRCVANYTG  418
                          *:********************.*:************************* ***.***

                                    Nidogen G2 beta-barrel
sp|P14543|NID1_HUMAN     NGRQCVAEGSPQRVNGKVKGRIFVGSSQVPIVFENTDLHSYVVMNHGRSYTAISTIPETV  480
sp|P10493|NID1_MOUSE     NGRQCVAEGSPQRVNGKVKGRIFVGSSQVPVVFENTDLHSYVVMNHGRSYTAISTIPETV  478
                          ****************************:****************************

sp|P14543|NID1_HUMAN     GYSLLPLAPVGGIIGWMFAVEQDGFKNGFSITGGEFTRQAEVTFVGHPGNLVIKQRFSGI  540
sp|P10493|NID1_MOUSE     GYSLLPLAPIGGIIGWMFAVEQDGFKNGFSITGGEFTRQAEVTFLGHPGKLVLKQQFSGI  538
                          *********:**********************************:****:**:**:****

sp|P14543|NID1_HUMAN     DEHGHLTIDTELEGRVPQIPFGSSVHIEPYTELYHYSTSVITSSSTREYTVTEPERDGAS  600
sp|P10493|NID1_MOUSE     DEHGHLTISTELEGRVPQIPYGASVHIEPYTELYHYSSSVITSSSTREYTVMEPDQDGAA  598
                          ********.***********:*:**************:************* **::***:

sp|P14543|NID1_HUMAN     PSRIYTYQWRQTITFQECVHDDSRPALPSTQQLSVDSVFVLYNQEEKILRYALSNSIGPV  660
sp|P10493|NID1_MOUSE     PSHTHIYQWRQTITFQECAHDDARPALPSTQQLSVDSVFVLYNKEERILRYALSNSIGPV  658
                          **: : ************.***:*********************:**:*************

sp|P14543|NID1_HUMAN     REGSPDALQNPCYIGTHGCDTNAACRPGPRTQFTCECSIGFRGDGRTCYDIDECSEQPSV  720
sp|P10493|NID1_MOUSE     RDGSPDALQNPCYIGTHGCDSNAACRPGPGTQFTCECSIGFRGDGQTCYDIDECSEQPSR  718
                          *:*****************:******** ***************:*************
```

Supplementary Figure 1: Structure and sequence of Nidogen G2 domain. A) Schematic representation of multi-domain human Nidogen 1 and its interactions with natural ligands. β-barrel structure within G2 domain is depicted. B) Secondary structure of Nidogen G2 β-barrel domain. In grey different β-sheets (A-K) and in blue the α-helix. C) Amino acid alignment between the human Nidogen 1 (P14543) and the mouse Nidogen 1 (P10493) proteins analyzed by Clustal Omega (EMBL-EMI). Black lines indicate the start and finish point of the G2 β-barrel domain (colored sequence). "*" indicates an amino acid match, ":" indicates amino acids with strong similar properties and "." indicates amino acids with weak similar properties.

GFP		HSNBT	CHANGE	BLOSUM62_SCORE	BLOSUM62_GROUP
ASN 149	→	THR 571	N – T	0	Same group
SER 202	→	SER 634	S – S	4	Identity
GLN 204	→	ASP 636	Q – D	0	Same group
ALA 206	→	SER 639	A – S	1	Same group
LEU 221	→	ALA 652	L – A	-1	Different group (both non-polar aliphatic)

Supplementary Figure 2: Comparative table of GFP residues predicted to be implicated in protein-protein contacts during protein assembly and amino acids present in the equivalent position within the HSNBT protein structure. Scores obtained by each substitution within block substitution matrix are indicated in the BLOSUM62_SCORE column. Positive scores indicate more likely substitution in a homologous sequence while negative values indicate less likely substitution in a homologous sequence.

Human Nidogen (P14543):

MLASSSRIRAAWTRALLLPLLLAGPVGCLSRQELFPFGPGQGDLELEDGDDFVSPALELSGALRFYDRSDIDAVYVTTNGIIATSEPPAKESHPGLFPPTFG
AVAPFLADLDTTDGLGKVYYREDLSPSITQRAAECVHRGFPEISFQPSSAVVVTWESVAPYQGPSRDPDQKGKRNTFQAVLASSDSSSYAIFLYPEDGLQ
FHTTFSKKENNQVPAVVAFSQGSVGFLWKSNGAYNIFANDRESVENLAKSSNSGQQGVWVFEIGSPATTNGVVPADVILGTEDGAEYDDEDEDYDLAT
TRLGLEDVGTTPFSYKALRRGGADTYSVPSVLSPRRAATERPLGPPTERTRSFQLAVETFHQQHPQVIDVDEVEETGVVFSYNTDSRQTCANNRHQCSV
HAECRDYATGFCCSCVAGYTGNGRQCVAEGSPQRVNGKVKGRIFVGSSQVPIVFENTDLHSYVVMNHGRSYTAISTIPETVGYSLLPLAPVGGIIGWMF
AVEQDGFKNGFSITGGEFTRQAEVTFVGHPGNLVIKQRFSGIDEHGHLTIDTELEGRVPQIPFGSSVHIEPYTELYHYSTSVITSSSTREYTVTEPERDGASP
SRIYTYQWRQTITFQECVHDDSRPALPSTQQLSVDSVFVLYNQEEKILRYALSNSIGPVREGSPDALQNPCYIGTHGCDTNAACRPGPRTQFTCECSIGFR
GDGRTCYDIDECSEQPSVCGSHTICNNHPGTFRCECVEGYQFSDEGTCVAVVDQRPINYCETGLHNCDIPQRAQCIYTGGSSYTCSCLPGFSGDGQACQ
DVDECQPSRCHPDAFCYNTPGSFTCQCKPGYQGDGFRCVPGEVEKTRCQHEREHILGAAGATDPQRPIPPGLFVPECDAHGHYAPTQCHGSTGYCWC
VDRDGREVEGTRTRPGMTPPCLSTVAPPIHQGPAVPTAVIPLPPGTHLLFAQTGKIERLPLEGNTMRKTEAKAFLHVPAKVIIGLAFDCVDKMVYWTDIT
EPSIGRASLHGGEPTTIIRQDLGSPEGIAVDHLGRNIFWTDSNLDRIEVAKLDGTQRRVLFETDLVNPRGIVTDSVRGNLYWTDWNRDNPKIETSYMDG
TNRRILVQDDLGLPNGLTFDAFSSQLCWVDAGTNRAECLNPSQPSRRKALEGLQYPFAVTSYGKNLYFTDWKMNSVVALDLAISKETDAFQPHKQTRLY
GITTALSQCPQGHNYCSVNNGGCTHLCLATPGSRTCRCPDNTLGVDCIEQK

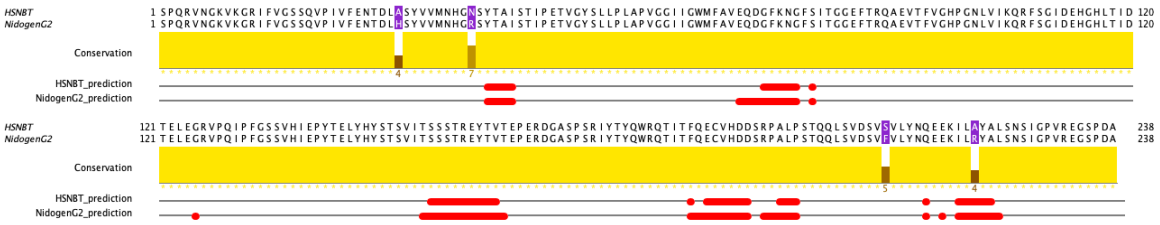
HSNBT:

SPQRVNGKVKGRIFVGSSQVPIVFENTDLASYVVMNHGNSYTAISTIPETVGYSLLPLAPVGGIIGWMFAVEQDGFKNGFSITGGEFTRQAEVTFVGHP
GNLVIKQRFSGIDEHGHLTIDTELEGRVPQIPFGSSVHIEPYTELYHYSTSVITSSSTREYTVTEPERDGASPSRIYTYQWRQTITFQECVHDDSRPALPSTQ
QLSVDSVSVLYNQEEKILAYALSNSIGPVREGSPDA

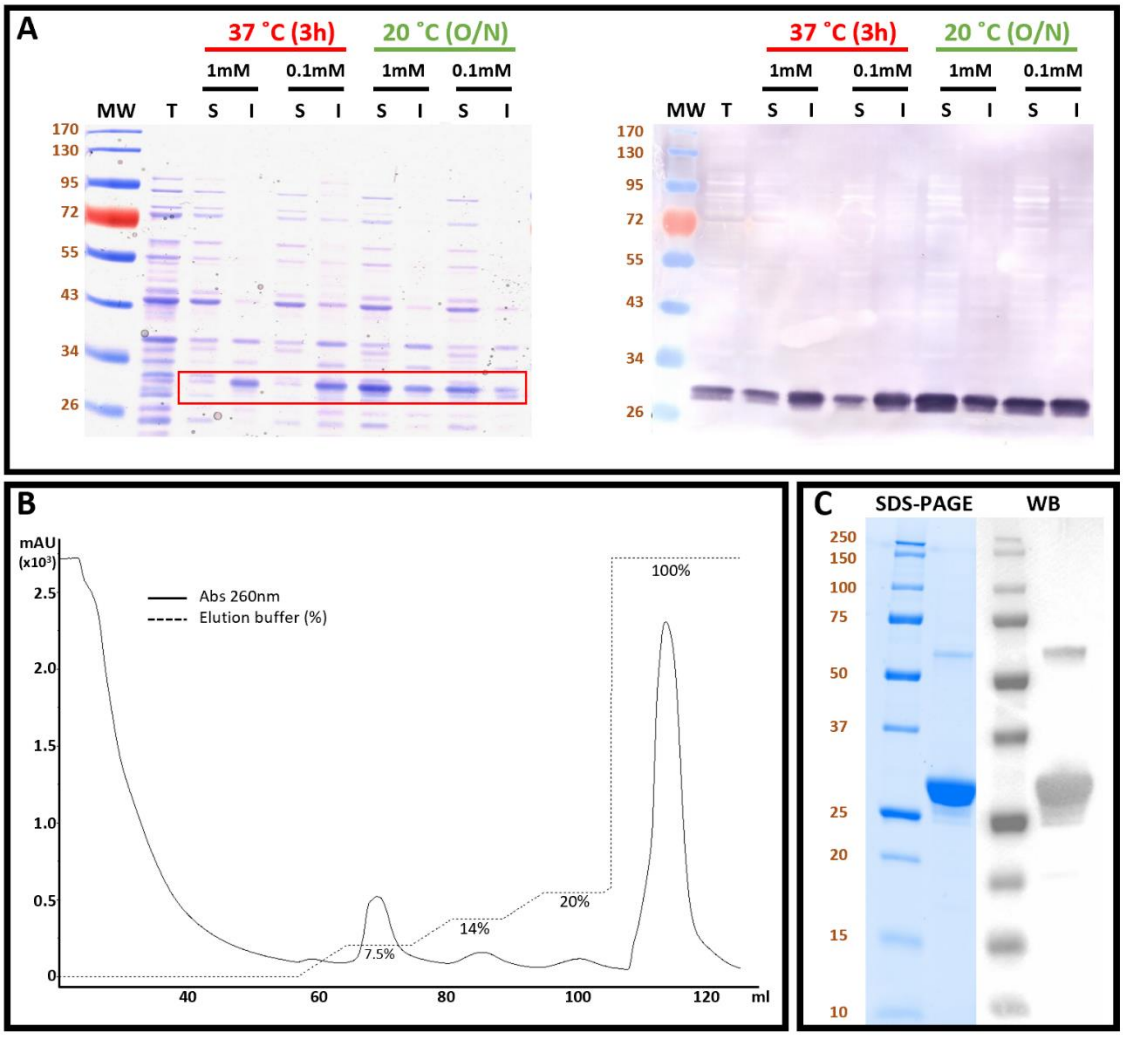
T22-HSNBT-H6:

RRWCYRKCYKGYCYRKCRGGSSRSSSPQRVNGKVKGRIFVGSSQVPIVFENTDLASYVVMNHGNSYTAISTIPETVGYSLLPLAPVGGIIGWMFAVEQ
DGFKNGFSITGGEFTRQAEVTFVGHPGNLVIKQRFSGIDEHGHLTIDTELEGRVPQIPFGSSVHIEPYTELYHYSTSVITSSSTREYTVTEPERDGASPSRIY
TYQWRQTITFQECVHDDSRPALPSTQQLSVDSVSVLYNQEEKILAYALSNSIGPVREGSPDAKHHHHHH

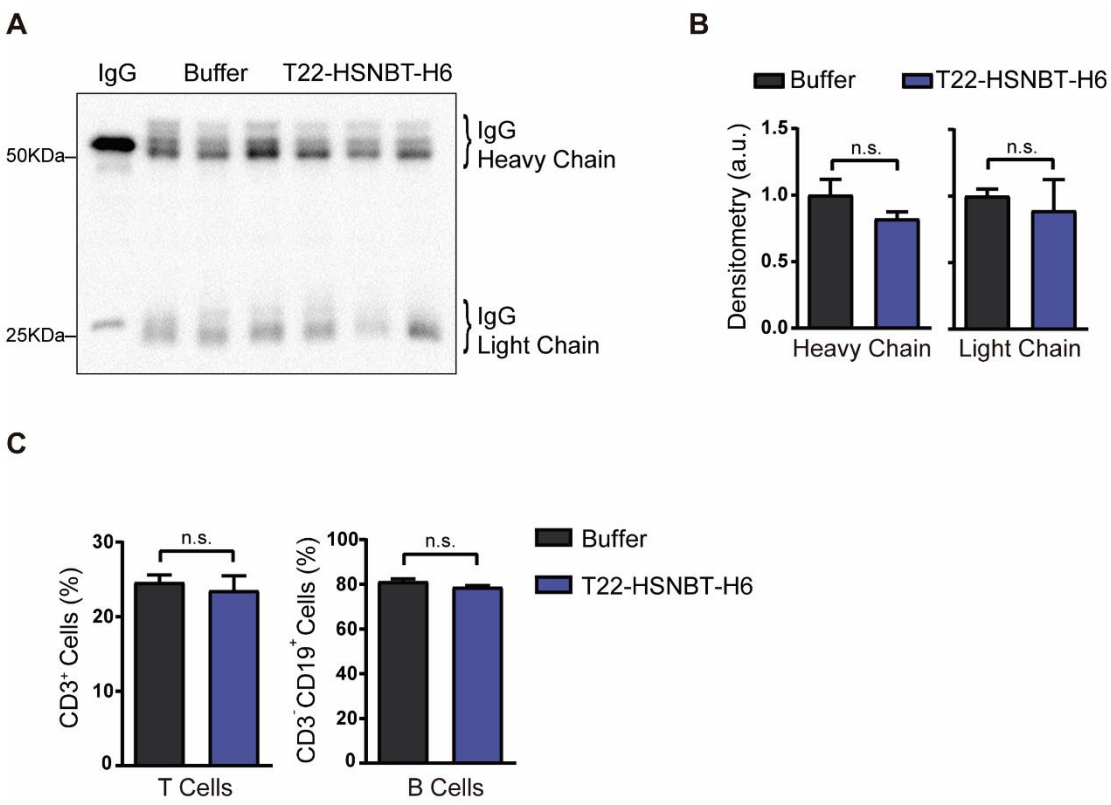
Supplementary Figure 3: Amino acid sequence of Human Nidogen 1, HSNBT scaffold and T22-HSNBT-H6 protein. In human Nidogen, G2 β-barrel domain is highlighted in blue and candidate amino acids to be mutated indicated in black. In HSNBT sequence, incorporated mutations are indicated in red. In T22-HSNBT-H6, N-terminal T22 ligand is indicated in purple, short linker is indicated in black, incorporated mutations are indicated in red and C-terminal poly-histidine tail is highlighted in dark blue.



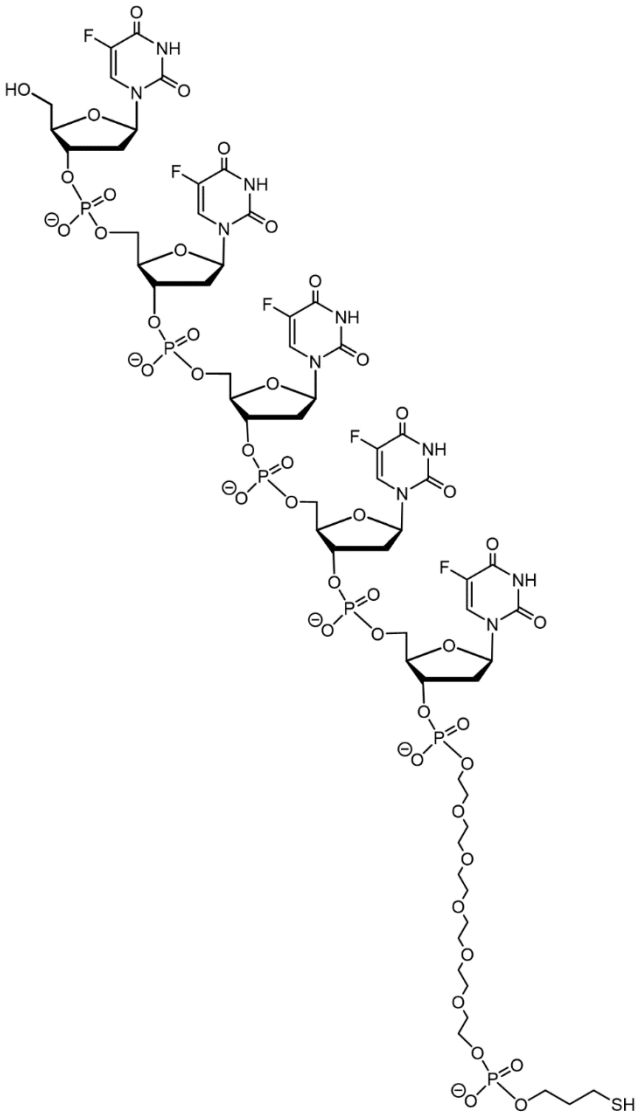
Supplementary Figure 4: Sequence alignment and *in silico* predictions for structural epitopes in human Nidogen G2 and engineered HSNBT proteins. Mutations in HSNBT are highlighted in violet. In the annotations below, the consensus prediction of structural epitopes is highlighted in red for each sequence.



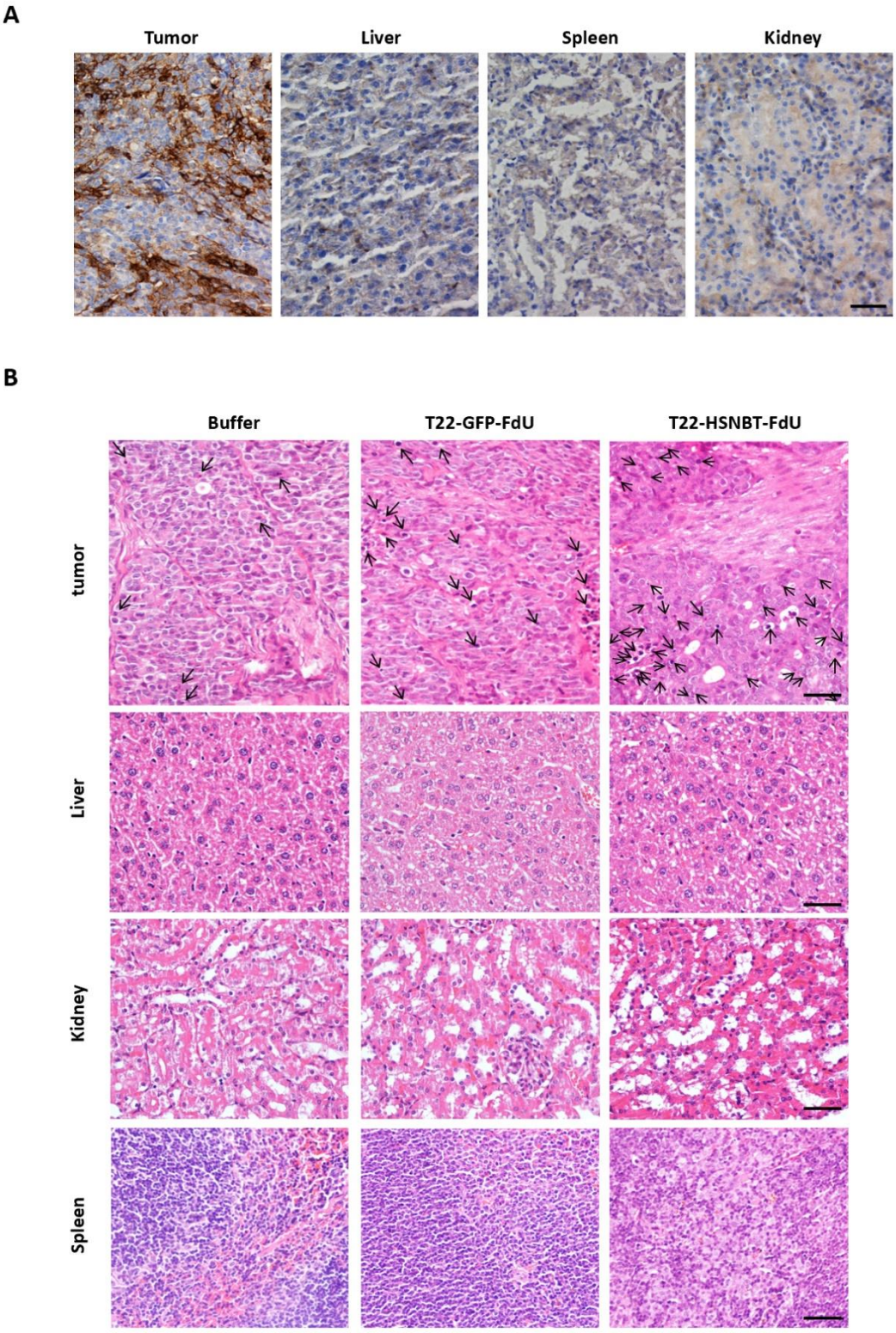
Supplementary Figure 5: Production and purification of T22-HSNBT-H6 protein. A) SDS-PAGE (left) and Western-blot (right) showing the production of T22-HSNBT-H6 protein in *E.coli* Origami B strain upon induction at different concentrations of IPTG (0.1 mM or 1 mM) and different temperatures (37ºC for 3h or 20ºC over night (O/N)). "T" indicates total cell fraction, "S" indicates soluble cell fraction and "I" indicates insoluble cell fraction. Red box in SDS-PAGE shows bands corresponding to T22-HSNBT-H6 protein. Protein in western-blot is immunodetected using an anti-His monoclonal antibody (Santa Cruz Biotechnology). B) IMAC purification chromatogram of T22-HSNBT-H6 protein. Black line indicate UV light absorbance at 260nm and punctuated line indicates the concentration of imidazole containing elution buffer. C) SDS-PAGE (left) and Western-blot (right) of purified T22-HSNBT-H6 protein showing pure protein monomer (30.3 kDa) and dimer (60.6kDa). Protein in western-blot is immunodetected using an anti-His monoclonal antibody (Santa Cruz Biotechnology).



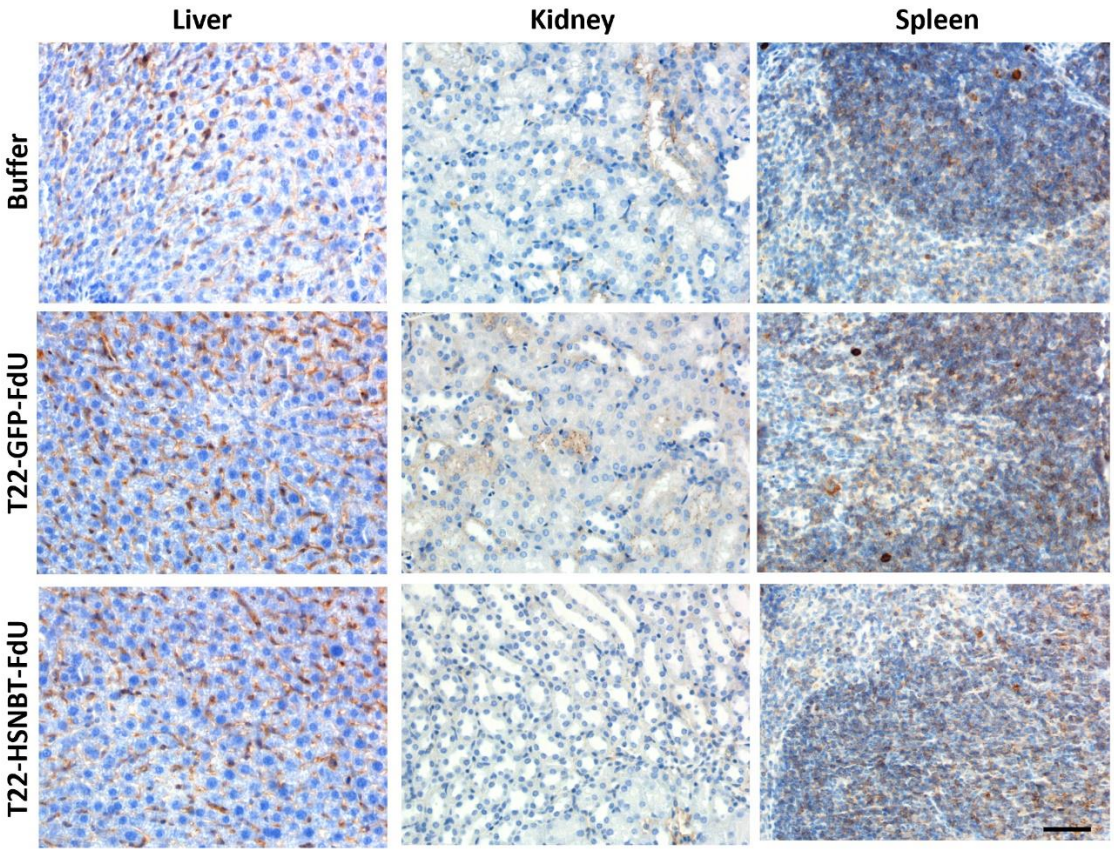
Supplementary Figure 6: Immune response in BALB/c mice after T22-HSNBT-H6 treatment. BALB/c mice were treated twice with buffer (100 mL) or 100 µg of T22-HSNBT-H6 intravenously (n=3). A) IgG present in mouse serum was detected by western blot (heavy chain upper band and light chain lower band). IgG positive control correspond to a purified anti-E-cadherin IgG (first lane). B) Densitometry quantification of heavy and light IgG chains. Values are represented as a fold change respect to the mean of buffer treated animals. C) Flow cytometry of mice spleen. Graphs shown quantification of both T Cell and B Cell in buffer and T22-HSNBT-H6 treated animals. T Cells are represented as the percentage of CD3 positive cells from all alive cells. B Cells are the percentage of CD19 positive cells from CD3 negative population. Data is presented as mean ± standard error. n.s. non statistically significant.



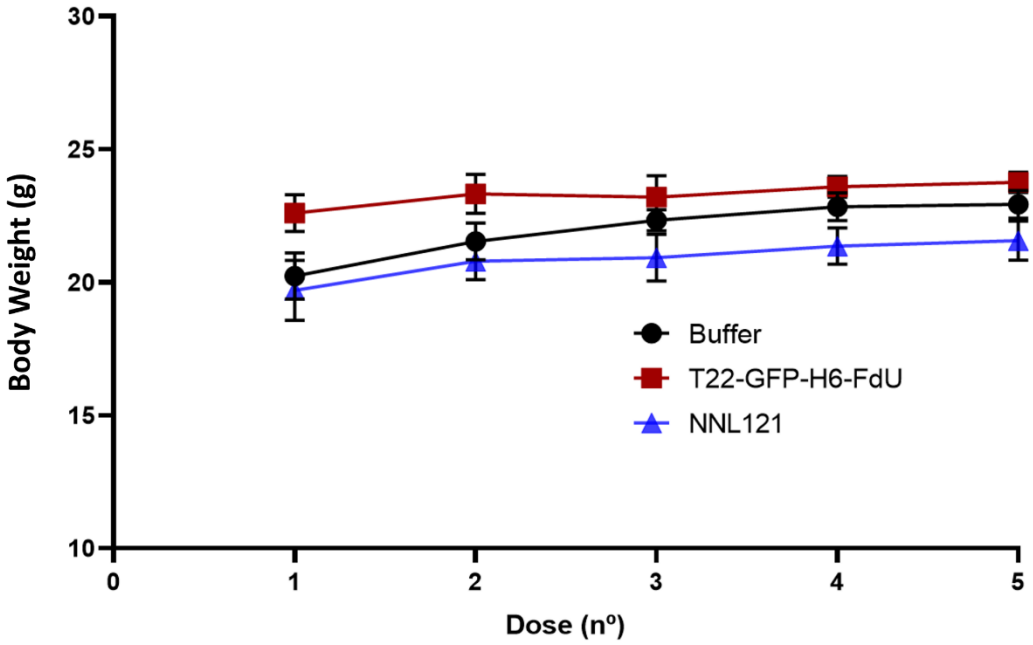
Supplementary Figure 7: Structure and composition of the 5-Fluoro-2'-deoxyuridine pentamer (oligo-FdU) carrying an hexaethyleneglycol (HEG) and a mercaptopropyl phosphate linkers.



Supplementary Figure 8: CXCR4 receptor expression and histological analysis of tumor and normal tissue slices. A) Anti-CXCR4 IHC staining of subcutaneous tumor and normal tissue slices (liver, kidney and spleen) showing CXCR4 overexpression exclusively in tumor tissue. B) Representative H&E staining of subcutaneous tumors and normal tissue slices (liver, kidney and spleen) showing apoptotic figures induced by each tested nanoconjugates (black arrows) 24h after a single dose administration, only in tumor tissue. Scale bars indicate 50 µm.



Supplementary Figure 9: Representative IHC images of caspase 3 activation in normal tissue (liver, kidney and spleen) 24h after single dose nanoconjugate administration showing no significant differences in apoptosis activation in any tissue between experimental groups or between the experimental and buffer treated groups. Scale bars indicate 50 μm.



Supplementary Figure 10: Mouse body weight variation. Measurement of total body weight of mice during repeated dose administration of buffer, T22-GFP-H6-FdU or T22-HSNBT-H6-FdU nanoconjugates (20 μg, three times a week, × 5 doses). Data is expressed as mean body weight ± standard error.

Soluble ligand	Immobilized Collagen type IV	Immobilized Perlecan
NidoG2-H6	2.13± 0.68	1.48 ± 0.36
T22-HSNBT-H6	n.d.	n.d.
GFP-H6	n.d.	n.d.

Supplementary Table 1: Binding of T22-HSNBT-H6 nanoparticles to immobilized Perlecan and Collagen type IV. Kd in μM obtained in surface plasmon resonance binding assays using soluble ligands at 70–10,000 nM. Values are shown as mean ± standard error. (n.d.): no binding detected up to 20μM ligand. Wild type Nidogen G2 domain and GFP proteins have been used as a positive and negative controls respectively.

References

[1] Marti-Renom MA, Stuart AC, Fiser A, Sanchez R, Melo F, Sali A. Comparative protein structure modeling of genes and genomes. Annual review of biophysics and biomolecular structure 2000;29:291-325.
[2] Rueda F, Cespedes MV, Conchillo-Sole O, Sanchez-Chardi A, Seras-Franzoso J, Cubarsi R, Gallardo A, Pesarrodona M, Ferrer-Miralles N, Daura X, Vazquez E, Garcia-Fruitos E, Mangues R, Unzueta U, Villaverde A. Bottom-Up Instructive Quality Control in the Biofabrication of Smart Protein Materials. Advanced materials 2015;27:7816-22.
[3] Gourlay LJ, Peri C, Ferrer-Navarro M, Conchillo-Sole O, Gori A, Rinchai D, Thomas RJ, Champion OL, Michell SL, Kewcharoenwong C, Nithichanon A, Lassaux P, Perletti L, Longhi R, Lertmemongkolchai G, Tilball RW, Daura X, Colombo G, Bolognesi M. Exploiting the Burkholderia pseudomallei acute phase antigen BPSL2765 for structure-based epitope discovery/design in structural vaccinology. Chemistry & biology 2013;20:1147-56.
[4] Lassaux P, Peri C, Ferrer-Navarro M, Gourlay LJ, Gori A, Conchillo-Sole O, Rinchai D, Lertmemongkolchai G, Longhi R, Daura X, Colombo G, Bolognesi M. A structure-based strategy for epitope discovery in Burkholderia pseudomallei OppA antigen. Structure 2013;21:167-75.
[5] Gaudesi D, Peri C, Quilici G, Gori A, Ferrer-Navarro M, Conchillo-Sole O, Thomas R, Nithichanon A, Lertmemongkolchai G, Titball R, Daura X, Colombo G, Musco G. Structure-based design of a B cell antigen from B. pseudomallei. ACS chemical biology 2015;10:803-12.
[6] Gourlay LJ, Thomas RJ, Peri C, Conchillo-Sole O, Ferrer-Navarro M, Nithichanon A, Vila J, Daura X, Lertmemongkolchai G, Titball R, Colombo G, Bolognesi M. From crystal structure to in silico epitope discovery in the Burkholderia pseudomallei flagellar hook-associated protein FlgK. The FEBS journal 2015;282:1319-33.
[7] Scarabelli G, Morra G, Colombo G. Predicting interaction sites from the energetics of isolated proteins: a new approach to epitope mapping. Biophysical journal 2010;98:1966-75.
[8] Peri C, Sole OC, Corrada D, Gori A, Daura X, Colombo G. Prediction of Antigenic B and T Cell Epitopes via Energy Decomposition Analysis: Description of the Web-Based Prediction Tool BEPPE. Methods in molecular biology 2015;1348:13-22.
[9] Fiorucci S, Zacharias M. Prediction of protein-protein interaction sites using electrostatic desolvation profiles. Biophysical journal 2010;98:1921-30.
[10] Shen MY, Sali A. Statistical potential for assessment and prediction of protein structures. Protein Sci 2006;15:2507-24.
[11] Schymkowitz J, Borg J, Stricher F, Nys R, Rousseau F, Serrano L. The FoldX web server: an online force field. Nucleic acids research 2005;33:W382-8.
[12] Bayne LJ, Vonderheide RH. Multicolor Flow Cytometric Analysis of Immune Cell Subsets in Tumor-Bearing Mice. Cold Spring Harbor protocols 2013;2013:pdb.prot077198.