Computational design of N-linked glycans for high throughput epitope profiling

Per Jr. Greisen*[1], Li Yi[2], Rong Zhou*[3], Jian Zhou[3], Eva Johansson[2], Tiantang Dong[3], Haimo Liu[3], Laust B Johnsen[2], Søren Lund[2], L. Anders Svensson[2], Haisun Zhu[3], Nidhin Thomas[1], Zhiru Yang[3], Henrik Østergaard*[4]

[1]Data Science and Innovation, [2]Global Research Technologies, [3]Discovery Technology China, [4]Global Drug Discovery, Novo Nordisk A/S

*Corresponding authors:

Per Jr. Greisen
pjug@novonordisk.com
Digital Science and Innovation, Novo Nordisk A/S

Rong Zhou
rzhou01@foxmail.com
Discovery Technology China, Novo Nordisk A/S

Henrik Østergaard
henrik3601@gmail.com
Global Drug Discovery, Novo Nordisk A/S

Abstract

Efficient identification of epitopes is crucial for drug discovery and design as it enables the selection of optimal epitopes, expansion of lead antibody diversity, and verification of binding interface. Although high resolution low throughput methods like X-ray crystallography can determine epitopes or protein-protein interactions accurately, they are time-consuming and can only be applied to a limited number of complexes. To overcome these limitations, we have developed a rapid computational method that incorporates N-linked glycans to mask epitopes or protein interaction surfaces, thereby providing an atomistic mapping of these regions. Using human coagulation factor IXa (fIXa) as a model system, we could rapidly and reliably delineate epitopes through the insertion of N-linked glycans that efficiently disrupt binding in a site-selective manner. To validate the efficacy of our method, we conducted ELISA experiments and high-throughput yeast

surface display assays. Furthermore, X-ray crystallography was employed to verify the results, thereby recapitulating through the method of N-linked glycans an atomistic resolution mapping of the epitope.

Introduction

The ability to rapidly identify epitopes and protein-protein interactions (PPI) is of utmost importance not only in interrogating new biology (1) but also in drug design, where identification and selection of optimal epitopes represent a key element in the optimization of potency and safety (2). It is important to have geometrically optimal epitopes to achieve the optimal potency of multispecific antibodies. This is particularly important in the case of bispecific antibodies like Mim8 (3) and emicizumab (4). Moreover, in *de novo* computational protein design, it is essential to rapidly determine whether the design is binding to the desired epitope. Here, the binding is often tested through methods like yeast surface display (YSD). Hence, there is a need for a fast method to structurally pinpoint the position of the epitope or determine if the binding interface is in accordance with computational design.

Traditional methods for mapping epitopes have certain limitations associated with them. One commonly used method is binning monoclonal antibodies (mAbs) into different epitope bins, thereby grouping the mAbs into overlapping or similar epitopes (5). Using full-length mAbs can be misleading due to steric hindrance of the fragment crystallizable (Fc) region, which does not necessarily reflect the binding region of the epitope. Other common ways of determining epitopes include mutational scanning, peptide mapping, NMR spectroscopy (6, 7), X-ray crystallography(8, 9), and cryo-electron microscopy(10). However, these are considered low throughput methods that require considerable effort or time and can test only a limited number of complexes. The first two methods have low resolution, while X-ray and cryo-EM of the co-complex can provide atomistic details of the interaction. Another popular method that can be used to profile epitopes and protein-protein interactions is hydrogen-deuterium exchange mass spectrometry (HDX-MS), which has been used to accurately map interactions between antibodies and antigens(11). The technique can accurately identify epitopes, but it is important to be cautious when analyzing the results as significant changes upon binding of the overall structure of the molecule can make it hard to interpret. In proteins with allosteric sites it can be difficult to map the precise epitope using HDX-MS due to conformational changes upon binding (12).

N-linked glycosylation is a common post-translational modification that occurs in many human proteins (13). It expands the functional properties of proteins by increasing the physico-chemical properties, e.g., stability (14). N-linked glycosylation of proteins occurs through a consensus Asn-Xxx-Ser/Thr recognition motif. Here, the sidechain of the initial Asn represents the glycan acceptor site; the second position (Xxx) can be any amino acid except proline and is followed by either serine(S) or threonine(T) at the third position. The efficiency of a glycosylated site depends on many factors, including the local 3-dimensional structure around the consensus motif (15) and

surface accessibility. Introducing a N-linked glycan within an epitope will disrupt ligand binding through simple steric hindrance. Others have investigated the idea of using a N-linked glycan to mask or protect an epitope with great success. Lombana et al. (16) developed the glycosylation-engineered epitope mapping (GEM) method, which involves the insertion of N-linked glycans at specific sites on the surface of the antigen to conceal localized regions. The method calculates the solvent accessible surface area (SASA) of the N and S/T residues and determines if they are exposed. There is no structural modeling of the N-linked glycan, and the method has not incorporated design of the S/T residue at the third position, which can have different preferences depending on local protein structure.

As discussed, identifying epitopes in a high throughput manner with high precision is desirable to understand biology and engineer optimal drugs targeting specific epitopes. Hence, we have improved on existing computational methods to insert and design the local sequence of the target protein such that it can accommodate a N-linked glycan at scale. It was achieved through a combination of structural modeling and computational design techniques and can be fully automated. Further, it is possible to add N-linked glycans at multiple positions of the human factor IXa without altering its overall structure (17). Using this model system of factor IXa, we were able to validate the method in accordance with a crystal structure of the complex using low- and high-throughput experimental methods.

Methods

Generation of anti-FIXa antibodies

Antibodies specific for human FIXa were generated in Kymouse mice (Kymab Group Ltd) that comprise a human antibody repertoire. Kymouse mice were immunized with active-site inhibited FIXa purchased from Haematologic technologies Inc/Prolytics. After electrofusion, hybridomas were plated in microtiter plates and supernatants were screened by a FIXa ELISA. Hybridomas secreting antibodies specific for FIXa were propagated and their supernatants subsequently used for proteinA affinity purification. FIXa binding was confirmed on fortebio, and hybridoma hits were sequenced and recombinantly expressed in HEK293 cells using standard techniques.

N-linked glycan library production and characterization

The DNA fragment encoding the amino acid sequence of HPC4 tagged human factor IX (fIX) without the gamma-carboxyglutamic (GLA) domain was synthesized and cloned into the pJSV vector with CD33 signal peptide to replace the native signal peptide(18). 98 variants were created, each containing specific engineered mutations for N-linked glycosylation. The mutations introduced were an asparagine (N) residue followed by any amino acid residue except proline (X) and a serine

(S) or threonine (T) residue at different positions in the DNA sequence. Each of the 98 variants were individually cloned for further study or experimentation. HEK293 6E cells were used for transient production of these proteins following standard protocols for 293Fectin (Invitrogen). The cell cultures were harvested at 5 d post-transfection by centrifugation at 6,000 rpm for 15 min, and the supernatants were filtered using a 0.45μm filter for purification with anti-HPC4 sepharose 4FF affinity columns (anti-HPC4 antibody coupled to the CNBr activated sepharose 4FF resin, GE), equilibrated in buffer (20mM Tris-HCl, 100mM NaCl, 1mM $CaCl_2$, pH 7.4). The bound proteins were eluted with buffer (20mM Tris-HCl, 100mM NaCl, pH 7.4, 1mM EGTA) as final products. The variants were analyzed with SEC-UPLC, LC-MS and SDS-PAGE, which were also digested with human factor XIa(19, 20) and then analyzed with SDS-PAGE to identify if the N-linked glycosylation was inserted.

FACS analysis

The yeast surface display plasmids were constructed to evaluate the binding of single-chain variable fragments (scFv) to different factor IXa variants. The *AGA2* gene downstream of the *GAL1* promoter was fused with a four-part cassette encoding the VL, linker, VH, and FLAG tag fragments, forming an VL–Linker–VH-FLAG-Aga2 cassette. Two plasmids that contain (G4S)3 and (G4S)4 linker, respectively, were constructed.

The constructed plasmids were transformed into EBY100 cells (*URA*$^+$, *leu*$^-$, *trp*$^-$) followed by cultivation and induction. The cells bearing different constructs were labeled with anti-FLAG-iFluor 488 and anti-HPC4-iFluor 647 antibodies (GenScript, China) followed by detection using similar protocols published previously(21) with the NovoCyte flow cytometer (Agilent, USA). The iFluor 647 fluorescent intensity was detected with the APC channel, 670/30 nm band pass, and iFluor 488 fluorescent intensity was detected with the FITC channel, 530/30 nm band pass. The binding efficiencies against different fIXa variants were evaluated as: Binding efficiency = [Mean Fluorescency Intensity of iFluor 647] /[Mean Fluorescency Intensity of iFluor 488], in which Mean Fluorescency Intensity (M.F.I.) of iFluor 647 and M.F.I. of iFluor 488 represent the binding of fIXa variants and surface display of scFv, respectively.

ELISA of N-linked glycosylation library screening

An anti-HPC4 monoclonal mouse antibody was coated onto 96-well ELISA plates (9018, Corning) at 5 ug/ml, 100 ul/well. The plates were incubated at 4°C overnight. The plates were washed 3 times with PBS, blocked with 1% BSA for 3 hours. The plates were washed 3 times before adding human factor IX or human factor IX with N-linked glycosylation at 1ug/ml, 100 ul/well. 1 hours later, the plates were washed 3 times before adding anti-FIX antibodies to be assessed at 1ug/ml, 100 ul/well. The plates were washed 3 times again before adding anti-hIgG-HRP (Invitrogen, Cat # A18823) antibody for detection. 30 mins later the plates were washed followed by measuring ELISA signaling at an absorbance of OD450.

Computational design of N-linked glycans

The crystal structure of fIXa (PDB ID: 1RFN(22)) was used as a starting point for calculating the Define Secondary Structure of Proteins (DSSP)(23) using RosettaScripts (24, 25). A glycan model was generated by utilizing the crystal structure (PDB ID: 4Z7Q(26)) with the glycan alpha-D-mannopyranose-(1-3)-[alpha-D-mannopyranose-(1-6)]beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose (chain K in crystal structure) creating the parameter file used in Rosetta. Protein structures were analyzed and visualized using PyMOL(27). For detail and list of commands and scripts used please see Appendix.

Production of NN-2003Fab

The DNA fragments encoding the amino acid sequence of fab (NN-2003Fab) heavy chain and light chain were synthesized and cloned into the pJSV vector (generated in-house using a *CMV* promoter) with CD33 signal peptide to replace the native signal peptide, respectively, which then were co-transfected into HEK293 6E cells for expression together following standard protocols for 293Fectin (Invitrogen). The cell cultures were harvested at 5 d post-transfection. Cells were removed from the culture by centrifugation at 6,000 rpm for 15 min, and the supernatants were filtered using a 0.45μm filter. The supernatants were loaded to HiTrap Protein G HP (5ml, GE), equilibrated in PBS. The bound proteins were eluted with elution buffer (0.1 M glycine-HCl, pH 2.7), the eluate was neutralized with buffer (1M Tris, pH9.0), which was further polished with HiLoad Superdex 200 16/60 (GE), equilibrated and eluted with buffer (25mM HEPES, 150mM NaCl, pH7.4).

Crystal structure determination of NN-2003Fab in complex with des-(Gla-EGF1) fIXa

Crystals of NN-2003 Fab mixed in a 1:1 molar ratio with human EGR-chloromethylketone active-site inhibited des-(Gla-EGF1) fIXa from Cambridge ProteinWorks were grown using the sitting drop vapor diffusion technique at 18 °C. A protein solution of 100 nl 6.9 mg/ml complex in 20mM Tris-HCl, pH 7.4, 50mM NaCl, 2.5mM CaCl2 was mixed with 100 nl of 0.5 M ammonium sulphate, 0.1 M sodium citrate pH 5.6, 1.0 M lithium sulphate as precipitant and incubated over 60 µl precipitant. The crystal was cryo protected in a solution consisting of 0.375 M ammonium sulphate, 75 mM sodium citrate pH 5.6, 0.75 M lithium sulphate, 4 % glycerol, 4 % ethylene glycol, 4.5 % sucrose, and 1 % glucose prior to flash cooling in liquid nitrogen. Diffraction data were collected at 100K at the Swiss Light Source beamline X06DA (1.00000 Å wavelength) using a Pilatus2M pixel detector from Dectris. Autoindexing, integration and scaling of the data were performed with programs from the XDS package(28). The asymmetric unit contains one Fab:FIXa complexes as judged from Matthew's coefficient analysis. The structure was determined by molecular replacement with Phaser(29) as implemented in the program suite Phenix(30) with the crystal structure of a homologous Fab and FIXa from PDB entry 3KCG(31) as search models. The correct amino acid

sequence for the Fab was model built using COOT(32) and thereafter the structure was refined using steps of Phenix refinement(33) and manual rebuilding in COOT. Diffraction data and refinement statistics are found in Supplementary Table 1.

Results

Computation-Guided epitope mapping of factor IXa

To profile the epitopes recognized by a panel of antibodies raised against fIXa with high throughput and atomistic resolution, an N-linked glycan variant library of fIXa was generated using a computational workflow (see Figure 1A). The computational protocol was generated to design N-linked glycans for high-throughput epitope profiling with atomistic resolution (see Figure 1B). Firstly, the crystal structure of fIXa, along with a model glycan was used as the input. The glycan was structurally inserted and modeled using RosettaMatch (34) on positions that were considered loops according to the DSSP algorithm in RosettaScripts (24, 25), resulting in 158 positions. From a statistical analysis, an epitope is comprised on average of 9–22 amino acid residues (35), and loop regions are often the most frequent regions used to insert N-linked glycosylation in human proteins. From the computational screen, number of positions were identified, all of which were designed to obtain the N-linked glycosylation motif N-X-S/T (where X represents all amino acids except proline). To increase the quality of the designed N-linked glycans, the 3-dimensional structures were ranked using RosettaScripts to compute the cutoffs on the interface energies, < 0 REU, and geometrical restraints, < 10, based on the structural models. This resulted in 98 positions that, according to the computational design, could accommodate N-linked glycosylation and were chosen for experimental verification (see Supplementary Table 3 for a full list of sequences used in the study).

It was decided to use fIXa without the gamma-carboxyglutamic acid-rich (GLA) domain, as the epitopes of interest were assumed to be distal to the membrane. Ninety-nine variants, including the wild type (WT), were expressed in the mammalian cell line HEK293 to ensure that they were glycosylated. It was possible to purify 92 of the engineered variants which was done in 2 steps. The first batch was produced in 30 ml cell culture while for variants not expressing here were upscaled to 300 mL cell culture. This resulted in concentrations ranging of 0.08 to 2.25 mg/mL for the 92 variants which were used for further characterization. To determine whether the engineered variants had successfully incorporated an N-linked glycan, the level of N-glycosylation of each variant was estimated by separating them based on their mass differences using SDS-PAGE. This method was used to detect and compare the molecular weights of the different protein variants, confirming whether the N-linked glycan had been successfully introduced. Ninety-eight desGla-hfIX variants were designed with a single *de novo* N-linked glycan site, with 77 variants in the protease domain, 12 variants in the EGF2 domain, and 9 variants in the hinge region (see Supplementary Figure 1). SDS-PAGE analysis was suitable for characterizing the glycosylation level only for variants in the protease domain. SDS-PAGE analysis showed that 52 out of the 77 variants were fully or partially glycosylated (for a comprehensive description, see Supplementary Table 2). Since

the mass difference of the EGF domains was too small to be detected using SDS-PAGE, the EGF2 variants that were engineered with N-linked glycosylation were tested using LC-MS. Seven of the variants in the EGF2 domain were shown to have incorporated N-linked glycosylation (see Figure 2). Table 1 shows the range of the incorporation efficiencies from 0 to 100% where five of the engineered variants were unquantifiable.

ELISA assay mapping of factor IXa epitope

To map the epitope of the anti-fIXa NN-2003 full length monoclonal antibody (mAb), 56 purified variants were first tested in an ELISA assay along with the WT fIXa. Out of these, four N-linked glycan variants (see Table 2) showed reduced signals in the binding assay comparable to the negative control buffer (see Figure 3A). This indicates that the N-linked glycan in these variants interfered with binding and is part of the epitope. Loss of binding was defined as an ELISA signal equivalent to that of the buffer used. To verify if this was due to site-selective masking of the epitope on fIXa, a crystal structure of the co-complex was solved. From the co-complex structure, it was observed that the inserted N-linked glycans were indeed in the binding epitope (see Figure 3B), thus validating the reduced binding signal observed by the ELISA.

Epitope Mapping Using Yeast Display

The first step in raising binders or verifying *de novo*-designed proteins often involves relying on display technology, such as phage panning or yeast display. Compared to phage display, yeast surface display (YSD) can provide a more straightforward quantitative evaluation. In this study, we used two scFv versions of NN-2003 with different linkers, G4Sx3 and G4Sx4 (NN-2003-L1 and NN-2003-L2, respectively), for epitope profiling using YSD. Four N-linked glycosylated fIXa variants were chosen to evaluate the sensitivity of the method, with three N-linked variants in the epitope and one outside of the epitope. Out of the four variants that exhibited reduced binding signals, the three variants with N-linked glycans in the epitope exhibited a reduced binding signal that was comparable to background levels. However, in the case of NN-7505, a binding signal was still observed in a concentration-dependent manner using fIXa as bait. This is in accordance with the binding mode and epitope observed in the crystal structure between NN-2003 and fIXa (see Figure 4).

Discussion

Rapid identification and characterization of epitopes are of utmost importance to optimizing drugs so that they can interact optimally with their targets. Rapidly profiling epitopes with high resolution can, for example, aid in drug discovery to also ensure that lead compounds have diverse

interactions with their target, and lastly, validate *de novo*–designed proteins. The ability to rapidly and accurately map epitopes on protein structures is essential for accelerating the engineering and design of proteins, giving confidence in the epitope selection. Here, we have created a fast, reliable, and fine-grained method to insert N-linked glycans using *in silico* methods that map them directly onto a crystal structure or a structural model of the target protein with the necessary sequence modifications to incorporate the N-linked glycan. We used a model system that applies the human coagulation fIXa, which was chosen to verify the protocol. It was shown that the method was robust and could be applied in both low- and high-throughput screening assays, such as ELISA and YSD. In the process of inserting the N-linked glycans, a conservative approach was chosen to mainly focus on loops as defined computationally by DSSP. It has been shown that loops have a higher probability of N-linked glycosylation in human proteins (26) but N-linked glycans are also observed in other secondary elements such as sheets and helices. This method can be easily extended to other secondary elements like helices or sheets by modifying the computational method to include these elements. In this study, to achieve full coverage of the fIXa surface, N-linked glycans were introduced with a maximum spacing of 19 residues between each glycan. As this distance has been observed in other studies of epitopes to ensure coverage of the area, it ensures full coverage of the target protein and increases the likelihood of masking all potential epitopes, even on large proteins like fIXa. Linear epitopes can be in the range of 4-12 amino acids(36) and one might need to expand the frequency of N-linked glycans to other secondary elements in the protein of interest. We did observe that not all the engineered N-linked glycans were incorporated into fIXa. Not all designed positions can sustain the incorporation of an N-linked glycan, and this needs to be considered when applying the method. In a few cases, the site was also not 100% glycosylated, which could potentially make interpretation difficult if the glycosylation level was not properly analyzed. The N-linked glycosylation-dependent method is constrained by the requirement that the target proteins must be expressed in host systems that can perform N-linked glycosylation. Therefore, proteins expressed only in *Escherichia coli* are unsuitable for this method. However, this computational approach is adaptable to any chemical modifications, such as PEGylation, which may offer greater flexibility in some cases. The method's scope is generalizable to these modifications, making it a versatile technique.

In the ELISA experiment, variability of the signal was observed, indicating heterogeneity for some of the variants. The heterogeneity observed in the binding signal can also be attributed to the occupancy of N-linked glycans, which may result in a reduced binding signal, although this does not necessarily mean the complete removal of the signal. This means that even if an N-linked glycan is present on the protein surface and partially masking the epitope, the scFv may still be able to bind to the protein with reduced affinity. However, we were able to get 59/92 N-linked glycosylated variants with N-linked glycans inserted into fIXa. Using ELISA from the binding results, we were able to determine the right epitope according to the crystal structure with a clear signal of where the epitope was located. Four of the N-linked glycan variants showed a reduced signal, and according to the crystal structure, they were the only variants that should disrupt binding. The computational de-novo design of protein binders is often evaluated using yeast display, but confirmation of epitope binding can be a time limitation. It was not clear whether applying the epitope mapping using YSD would be as sensitive as the ELISA. However, the results showed that we were able to map the epitope using YSD by using four variants with engineered N-linked glycans close to the epitope. Three of the N-linked glycans showed reduced signals in ELISA, and a similar result was observed using titrations in YSD. The results were encouraging, as there were multiple nonspecific interactions between the engineering protein and its N-linked glycans that could be

anticipated. To conclude, we have developed an *in silico* method to insert N-linked glycans using structural modeling. The method was validated using fIXa as a model system, and successfully performed epitope mapping using experimental methods, such as ELISA and yeast surface display.

Figures

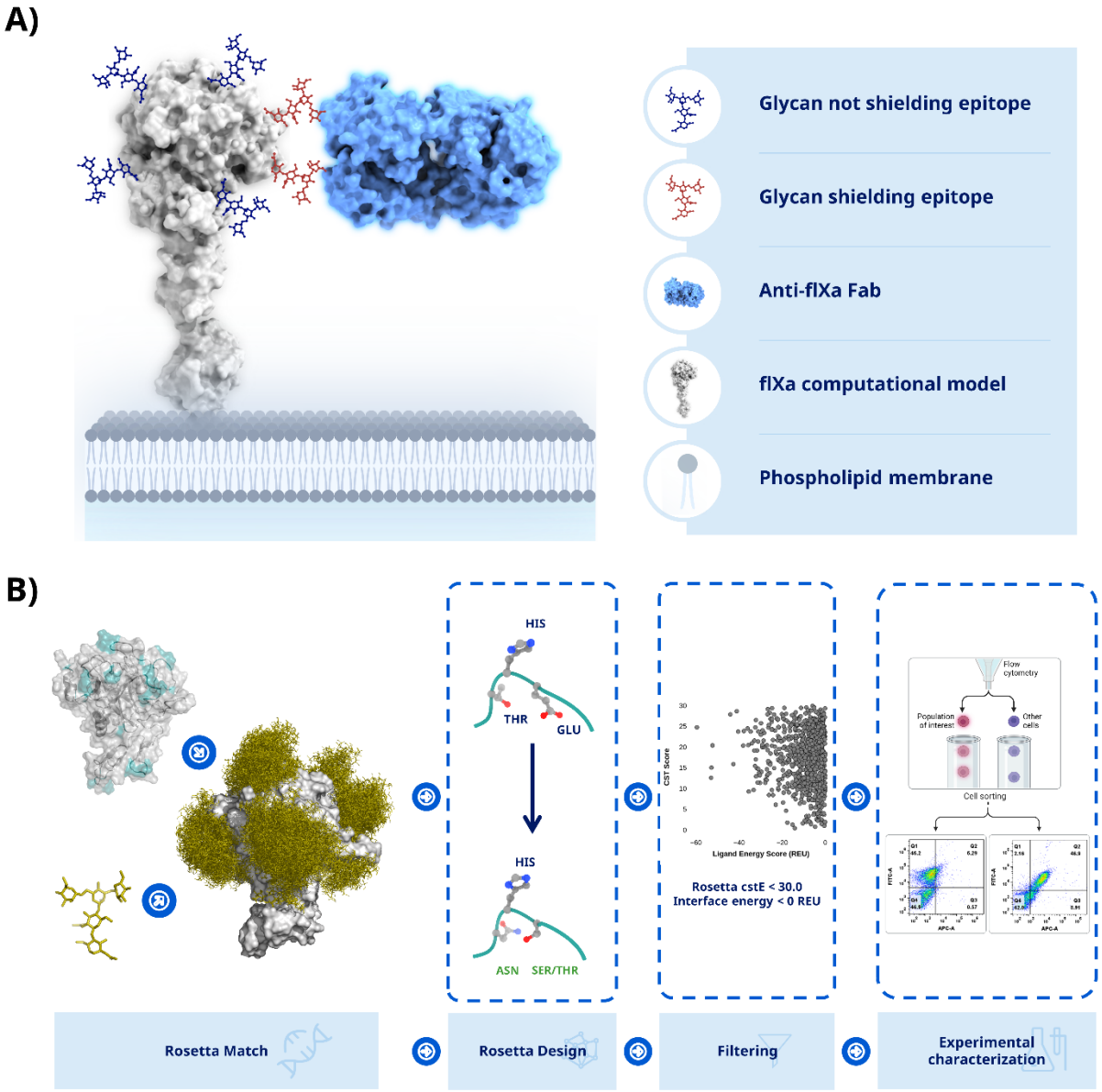


Figure 1 Epitope mapping through engineered N-linked glycans. A) The idea is to engineer N-linked glycans into the protein of interest which will mask the epitope/interface and thereby prevent binding/interaction. B) Firstly, loop elements(cyan) are identified using DSSP on a crystal structure of factor IXa which returns potential positions to place the N-linked glycan. Next, a model of glycan structure is generated which is used by RosettaMatch to insert the N-linked glycan onto the structure of fIXa. B) The sequence is designed by RosettaDesign to obtain the consensus motif N-X-S/T (X can be any amino acid except

proline) and optimized according to the input geometry. C) Engineered variants are filtered based on deviations from the specified geometry, < 10, and an interface energy < 0 REU is required. D) The remaining designs are used to experimentally characterize binding epitopes using either yeast surface display or other binding methods such as ELISA.

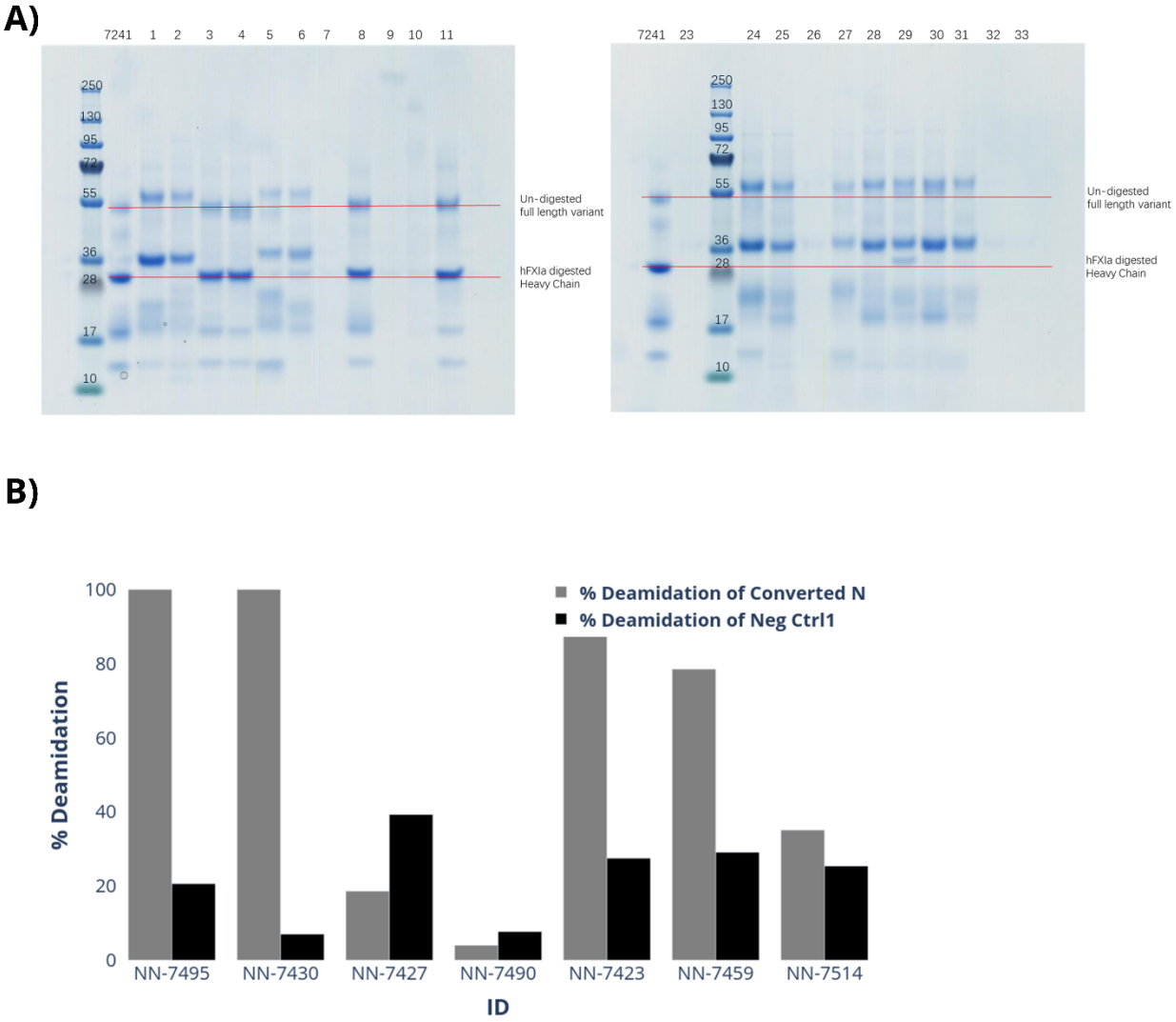


Figure 2 Characterization of N-linked glycosylation in EGF2 domain. A) The engineered variants are characterized on a reducing SDS-PAGE gel with NN-7241 (fIXa without gamma-carboxyglutamic acid domains) as reference. Shifts are indicated with respect to the heavy chain of fIXa (red line) B) The glycosylation level of engineered fIXa variants could be characterized by LC-MS/MS according to convert Asparagine to Aspartic Acid with removing N-linked glycosylation. To mitigate the false positive effect induced from chemical deamidation, deamidation ratio of N175 was calculated as a negative control. As an example, the LC-MS/MS analysis indicated that 3 of 7 variants were 100% incorporated with N-linked glycosylation; 1 variant was partially (78.5% or lower) incorporated with N-linked glycosylation and 3 variants was not or limited (35.1~4.0% or lower) incorporated with N-linked glycosylation.

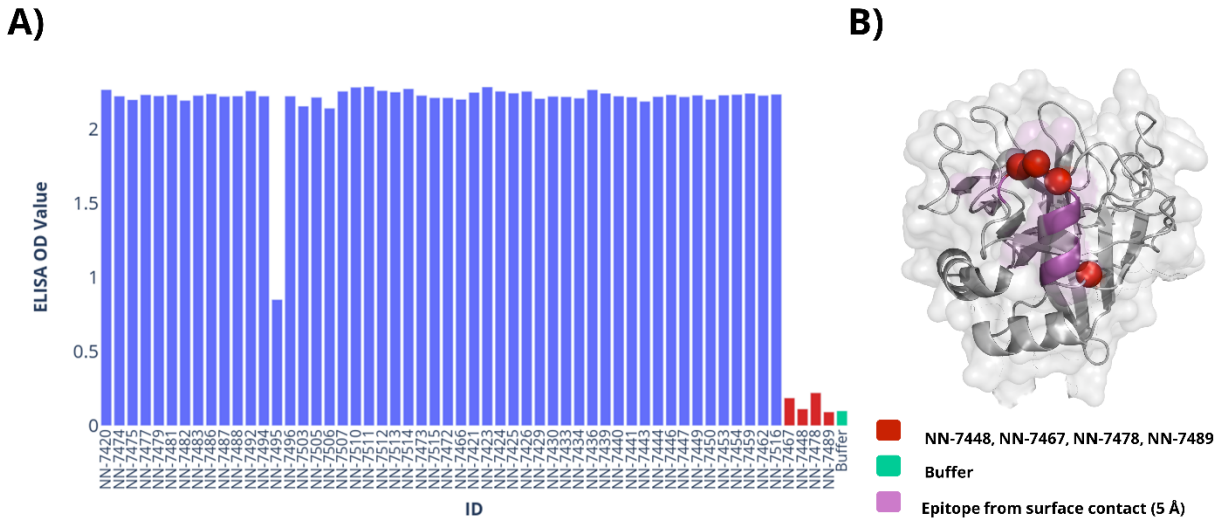


Figure 3 ELISA mapping of epitope. A) Measuring ELISA signaling for each of the N-linked glycosylated variants displaying the maximum value. Variants not showing reduced signal are in blue and the signal from the buffer (green) while variants with signal similar to the buffer are considered none binding (red). B) The 4 variants with reduced signal and their positions (red spheres for C-alpha position) relative to the epitope as defined by the crystal structure (purple).

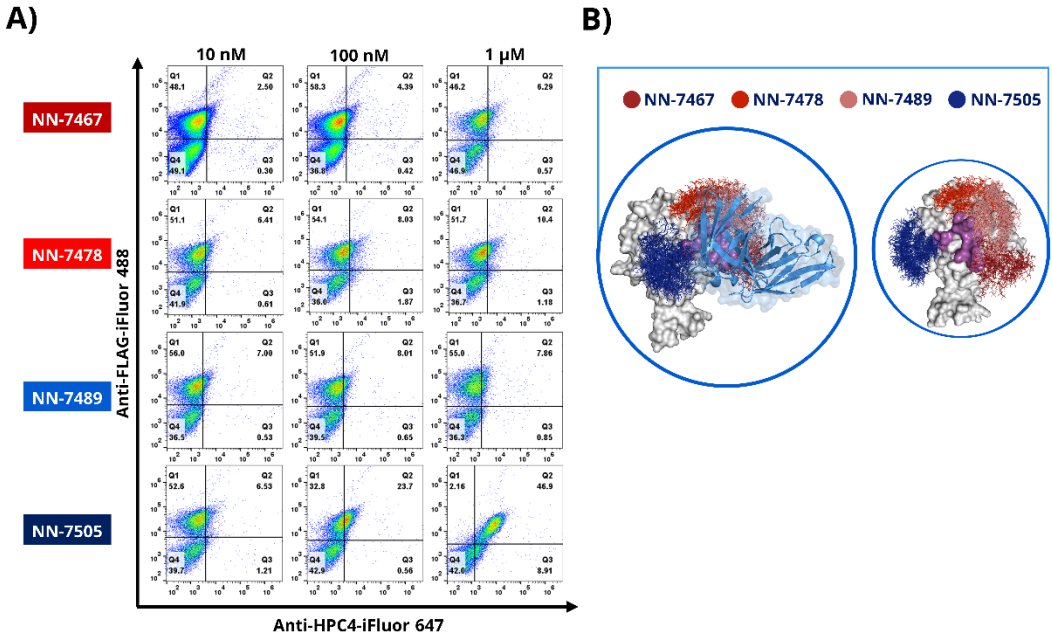


Figure 4 Epitope mapping by yeast display. A) Yeast surface displayed scFv NN-2003 (FLAG coupled with anti-FLAG-iFluor 488, y-axis) binding of 4 fIXa variants with N-linked glycans (coupled with anti-HPC4-iFluor 647, x-axis) on protein titrated at three different concentrations: 10 nM, 100 nM, and 1 uM. B) Crystal structure of the complex between NN-2003's fragment antigen-binding region (skyblue cartoon representation) and fIXa (white surface repsentation) with N-linked glycans

modeled onto the structure where red colors represent epitopes being masked by N-linked glycan while blue indicates no or little effect on binding.

Tables

Supplementary Table 1 Diffraction data and refinement statistics for the crystal structure determination of NN-2003Fab in complex with des-(Gla-EGF1) fIXa.

Wavelength (Å)	1.0000
Resolution range (Å)	49.98-2.602 (2.695-2.602)
Space group	I 2 2 2
Unit cell (Å, deg)	61.44 97.27 366.47 90 90 90
Total reflections	160379 (8632)
Unique reflections	33502 (2497)
Multiplicity	4.8 (3.5)
Completeness (%)	96.24 (73.60)
Mean I/sigma(I)	10.62 (1.28)
Wilson B-factor (Å^2)	50.02
R-merge	0.1322 (0.806)
R-meas	0.148 (0.9306)
R-pim	0.06544 (0.452)
CC1/2	0.996 (0.69)
CC*	0.999 (0.904)
Reflections used in refinement	33141 (2487)
Reflections used for R-free	1657 (124)
R-work	0.2670 (0.3629)
R-free	0.3236 (0.4278)
CC(work)	0.870 (0.672)
CC(free)	0.828 (0.538)
Number of non-hydrogen atoms	5672
macromolecules	5615
ligands	20
solvent	37
Protein residues	725
RMS(bonds)	0.003
RMS(angles)	0.56
Ramachandran favored (%)	93.28
Ramachandran allowed (%)	6.44
Ramachandran outliers (%)	0.28
Rotamer outliers (%)	3.22
Clashscore	5.76

Average B-factor (Å²)	61.15
macromolecules	61.22
ligands	74.69
solvent	43.23
Number of TLS groups	1
PDB entry	8OL9

Statistics for the highest-resolution shell are shown in parentheses.

Table 1 Quantification of the percentage of N-glycosylation presenting in EGF2 variants with peptide mapping.

NNCD	Serial Number	Glycosylation (%)	Mutations	Position
NN-7501	82	0	T66N G68S	EGF2
NN-7437	18	27	A72N N74S	EGF2
NN-7470	51	56.8	S56N D58S	EGF2
NN-7493	74	57.3	K60N V62S	EGF2
NN-7435	16	67.2	Q75N	EGF2
NN-7438	19	100	E73N Q75T	EGF2
NN-7460	41	100	A57N N59S	EGF2
NN-7427	8	ND	T41N N43S	EGF2
NN-7430	11	ND	D58N K60T	EGF2
NN-7456	37	ND	E67N Y69S	EGF2
NN-7469	50	ND	E50N F52S	EGF2
NN-7490	71	ND	S77N E79S	EGF2

ND means “not determined”.

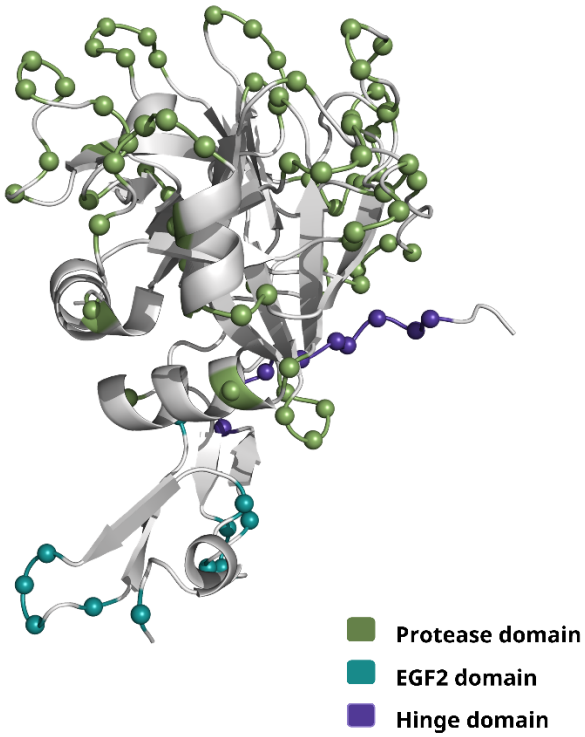
Table 2 N-linked glycosylated variants with reduced binding to NN-2003.

NNCD	Mutations	Epitope
NN-7448	F296N I298S	Yes
NN-7467	V285N R287T	Yes
NN-7478	T294N F296S	Yes
NN-7489	K295N	Yes

Acknowledgement

The authors thank the entire Novo Nordisk Mim8 project team for contributions, in particular, Xiaoli Yan for expression of these variants in the library, Xinping Qu, Lin Li, Chenxi Shen for purification of these variants, Xiaoai Wu for collecting data. We would like to express our gratitude to Bjarne Gram Hansen and Ida Hilden for enabling and guiding us in exploring scientific avenues.

Appendix



Supplementary Figure 1 Distribution of C-alpha positions of N-linked glycans into three domains of fIXa protease domain(smudge green spheres), hinge region (purpleblue spheres), and EGF2 domain (deepteal spheres). Figure was made with PyMOL(27).

Predict loop structure for matching.

The secondary structure was assign using dssp implemented in Rosetta:

rosetta_scripts.linuxgccrelease -database database -parser:protocol dssp.xml -in:file:s PDBFILE @flags

where the xml file, dssp.xml, was given as:

<ROSETTASCRIPTS>

<SCOREFXNS>

<enzdes weights=talaris2013 />

<soft weights=ligand_soft_rep />

</SCOREFXNS>

<TASKOPERATIONS>

<InitializeFromCommandline name=init/>

<SetIGType name=linmem_ig lin_mem_ig=1/>

<LimitAromaChi2 name=limchi2/>

</TASKOPERATIONS>

<FILTERS>

</FILTERS>

<MOVERS>

Minimization of complex - no design allowed

<TaskAwareMinMover name=min bb=0 chi=1 jump=1 scorefxn=enzdes task_operations=init/>

Packing of rotamers making sure no aromatic with chi2 of 90 degrees

<PackRotamersMover name=repack scorefxn=enzdes task_operations=init,limchi2/>

<ParsedProtocol name=min_repack_min>

<Add mover=min/>

<Add mover=repack/>

<Add mover=min/>

</ParsedProtocol>

<Dssp name=dssp reduced_IG_as_L=0/>

</MOVERS>

<PROTOCOLS>

<Add mover_name=dssp/>

</PROTOCOLS>

</ROSETTASCRIPTS>

using the following flags:

-ignore_zero_occupancy false

-packing

-use_input_sc

-ex1

-ex2

-ex1aro

-ex2aro

-linmem_ig 10

-no_optH false

-flip_HNQ

-no_his_his_pairE

-nblist_autoupdate

Match for N-glycan using Rosetta Matcher.

The geometrical search is performed using the executable below

match.linuxgccrelease -in:file:s PDBFILE -match:lig_name NAG -match:geometric_constraint_file

CSTFILE extra_res_fa NAG.fa.params -match:scaffold_active_site_residues POSITIONFILE

where the geometrical restraints have been specified for bond and angle:

CST::BEGIN

TEMPLATE:: ATOM_MAP: 1 atom_name: C1 O3 C2

TEMPLATE:: ATOM_MAP: 1 residue3: NAG

TEMPLATE:: ATOM_MAP: 2 atom_type: NH2O,

TEMPLATE:: ATOM_MAP: 2 residue1: N

CONSTRAINT:: distanceAB: 1.40 0.10 1000.00 1 0

CONSTRAINT:: angle_A: 105.00 1.00 100.00 360.00 1

CONSTRAINT:: angle_B: 120.00 5.00 100.00 360.00 1

CONSTRAINT:: torsion_A: -164.20 10.00 50.00 360.00 1

CONSTRAINT:: torsion_B: -120.00 10.00 50.00 360.00 1

CONSTRAINT:: torsion_AB: -117.00 360.00 0.00 360.00 12

CST::END

The following positions were screened:

1 3 4 7 8 9 10 11 12 13 14 20 21 22 23 24 25 33 34 39 43 44 45 46 47 48 49 54 55 56 58 59 60 61 62

63 64 65 76 77 78 79 81 82 83 85 86 87 88 89 90 96 97 98 99 100 101 103 104 106 107 108 110 111

112 123 130 131 133 134 135 136 137 138 140 141 143 151 152 160 161 162 163 164 165 166 167

172 173 174 175 176 177 178 180 181 182 183 184 185 186 187 193 194 195 196 207 208 209 210

211 212 213 214 215 224 233 234 235 236 237 238 239 240 246 247 248 253 254 255 256 257 262

263 264 265 269 270 271 272 273 274 278 279 280 281 282 283 284 285 286 287 288 289 290 291

292 293

Geometrical restraint score for each match.

To evaluate each match we score how well the geometrical parameters are obtained on the structure

rosetta_scripts.linuxgccrelease -database rosetta_database -parser:protocol score_pose.xml -in:file:s

PDBFILE -extra_res_fa NAG.fa.params @flags -parser_read_cloud_pdb 1

the flag file is defined the following way:

-preserve_header 1

-parser_read_cloud_pdb 1

-packing

-use_input_sc

-ex1

-ex2

-ex1aro

-ex2aro

-linmem_ig 10

-no_optH false

-flip_HNQ

-no_his_his_pairE

-nblist_autoupdate

-ignore_zero_occupancy false

and score_pose.xml is:

<ROSETTASCRIPTS>

<TASKOPERATIONS>

<InitializeFromCommandline name=init/>

<ReadResfile name=res filename=resfile />

<LimitAromaChi2 name=limchi2/>

</TASKOPERATIONS>

<SCOREFXNS>

<enzdes weights=talaris2013_cst />

</SCOREFXNS>

<MOVERS>

<AddOrRemoveMatchCsts name=addcst cstfile="nglycan.cst" cst_instruction=add_new/>

Minimization of complex - no design allowed

<TaskAwareMinMover name=min bb=0 chi=1 jump=1 scorefxn=enzdes task_operations=init/>

Packing of rotamers making sure no aromatic with chi2 of 90 degrees

<PackRotamersMover name=repack scorefxn=enzdes task_operations=init,limchi2,res/>

<ParsedProtocol name=min_repack_min>

<Add mover=min/>

<Add mover=repack/>

<Add mover=min/>

</ParsedProtocol>

</MOVERS>

<FILTERS>

<EnzScore name=cst_score scorefxn=enzdes whole_pose=1 score_type=cstE

energy_cutoff=10.0/>

</FILTERS>

<PROTOCOLS>

<Add mover=addcst/>

<Add mover=min_repack_min/>

<Add filter=cst_score/>

</PROTOCOLS>

<OUTPUT scorefxn=enzdes/>

</ROSETTASCRIPTS>

Python script to convert dssp output to position file.

To sample loop regions of the protein, dssp is used to determine loop regions in the protein. The dssp

output from Rosetta is converted into position file used by the match algorithm.

```
# Convert the DSSP output from Rosetta into a

# position file used for the match algorithm

#

# python convert_dssp_to_positionfile.py dsspfile

#

# generates the position file

# position.pos

import sys

inputfile = sys.argv[1]

tmp_pos = []

with open( inputfile ,'r' ) as f:

for line in f:

tmp_pos.append( line )

with open("position.pos",'w') as f:

for dssp in range(len (tmp_pos[0]) ):

if tmp_pos[0][dssp] == 'L' :

f.write( str(dssp+1) +" ")
```

Calculating DSSP of crystal structure (PDB ID: 1RFN).

The dssp is computated using Rosetta to assign secondary structure

LELLEELLLLLLLEEEEELLLLLEEEEEEELLEEEELHHHLLLLLLLEEEELLLELLLLLLLLEE

EEEEEEEELLLLELLLELLLLLEEEEELLLLLLELLELLLELLLHHHHHHHHHHLEEEEEELLEL

LLLLLELLELEEEEEEELLHHHHHHHLLLLLLLEEEELLLLLLELLLLLLLEEEEELLLEEE

EEEEEEELLLLLLLLEEEEEHHHLHHHHHHHHLLLLLLLHHHHHLLEEEELLLLEEEELLL

LEEELLLLLEEELLLLLLLLLLLLLL

Characterization of N-linked variants

ID	Serial Number	Conc. (mg/ml)	% Purity	Glycosylation	Mutations
NN-7420	1	0.56	59.72	Yes	E178N G180T
NN-7421	2	0.36	43.26	Yes	A215N I217T
NN-7422	3	0.43	64.54	Yes (Partial) >50%	Q93N
NN-7423	4	0.43	65.45	Yes	S92N
NN-7424	5	0.41	56.49	Yes	E194N E196S
NN-7425	6	0.34	45.56	Yes	K348N G350S
NN-7426	7	1.33	67.62	Yes	V181N I183T
NN-7427	8	0.46	64.31	No	T41N N43S
NN-7428	9	1.19	31.66	ND	R88N
NN-7429	10	0.46	43.84	Yes	G144N F146S
NN-7430	11	0.50	64.67	Yes	D58N K60T
NN-7431	12	0.14	39.48	No	V234N N236S
NN-7432	13	0.00	ND	ND	Q145N P147S
NN-7433	14	0.31	52.88	Yes	I192N E194T
NN-7434	15	0.40	55.04	Yes (Partial) <50%	M345N G347S
NN-7435	16	0.45	62.08	Yes (Partial) >50%	Q75N
NN-7436	17	0.63	65.22	Yes	K155N D157S
NN-7437	18	0.27	42.47	Yes (Partial) <50%	A72N N74S
NN-7438	19	0.34	50.85	Yes	E73N Q75T
NN-7439	20	0.48	62.66	Yes	G137N E139S
NN-7440	21	0.77	70.13	Yes	V156N A158S
NN-7441	22	0.12	37.28	Yes	D230N P232S
NN-7442	23	0.48	24.17	ND	Y213N A215S

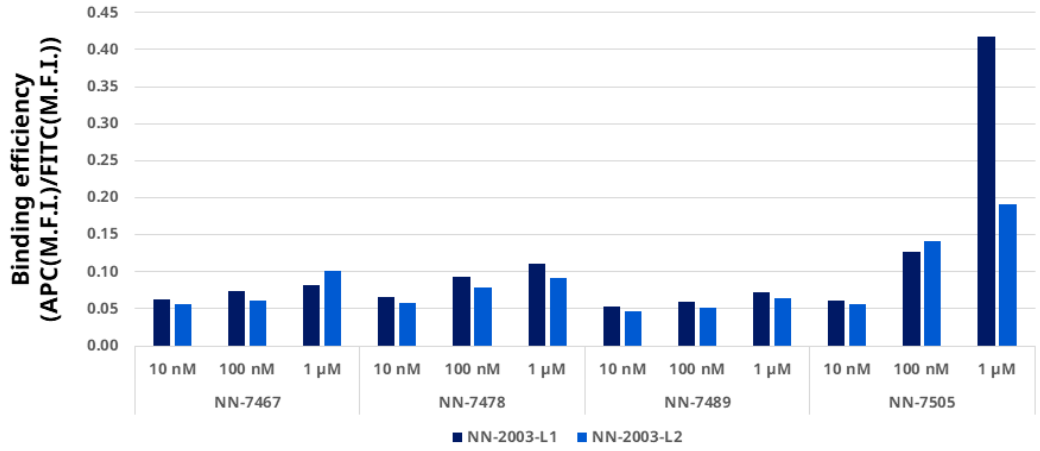
NN-7443	24	0.56	56.25	Yes	F159N G161S
NN-7444	25	0.36	44.93	Yes	F307N E309T
NN-7445	26	0.40	19.72	ND	H211N Y213S
NN-7446	27	0.22	50.96	Yes	P143N Q145S
NN-7447	28	0.42	50.32	Yes	D286N A288T
NN-7448	29	0.38	57.98	Yes	F296N I298S
NN-7449	30	0.51	58.67	Yes	E309N G311S
NN-7450	31	0.34	61.21	Yes	K346N K348T
NN-7451	32	0.46	30.76	ND	L235N
NN-7452	33	0.00	ND	ND	Y220N H222S
NN-7453	34	0.62	62.11	Yes (Partial)<50%	V267N H269S
NN-7454	35	0.33	41.68	Yes	D246N E248T
NN-7455	36	0.20	40.03	ND	T198N Q200S
NN-7456	37	0.36	43.02	ND	E67N Y69S
NN-7457	38	0.08	30.45	Yes	P83N P85S
NN-7458	39	0.10	36.19	ND	E199N K201S
NN-7459	40	0.56	59.52	Yes	H308N G310S
NN-7460	41	0.39	58.54	Yes	A57N N59S
NN-7461	42	0.24	47.99	Yes (Partial) >50%	A81N P83S
NN-7462	43	0.82	68.43	Yes	K182N
NN-7463	44	0.00	ND	ND	E231N L233T
NN-7464	45	0.30	53.67	No	G317N
NN-7465	46	0.11	32.48	ND	K142N G144S
NN-7466	47	0.41	60.52	Yes (Partial) <50%	R312N
NN-7467	48	1.63	56.54	ND	V285N R287T
NN-7468	49	0.24	49.51	ND	H190N I192S
NN-7469	50	0.53	22.55	ND	E50N F52S
NN-7470	51	0.35	50.09	Yes (Partial) >50%	S56N D58S
NN-7471	52	0.00	ND	ND	E167N W169S
NN-7472	53	0.49	61.45	Yes	H197N E199T

NN-7473	54	0.62	67.27	Yes	A216N N218S
NN-7474	55	0.58	69.75	Yes	T179N V181T
NN-7475	56	0.65	64.04	Yes	E193N
NN-7476	57	1.81	26.78	ND	V327N G329S
NN-7477	58	0.59	54.02	Yes	L275N L277S
NN-7478	59	0.48	59.31	Yes	T294N F296S
NN-7479	60	0.66	72.26	Yes	G257N G259S
NN-7480	61	0.30	50.37	No	K367N
NN-7481	62	0.58	64.07	Yes	K270N R272T
NN-7482	63	0.29	52.92	Yes (Partial) >50%	E196N
NN-7483	64	0.79	66.25	Yes	H269N G271S
NN-7484	65	0.46	58.52	No	T195N H197S
NN-7485	66	0.20	42.28	No	G329N
NN-7486	67	0.60	58.48	Yes	D157N F159S
NN-7487	68	0.14	35.11	Yes	K219N N221S
NN-7488	69	0.66	64.59	Yes	G180N K182T
NN-7489	70	0.80	78.28	Yes	K295N
NN-7490	71	0.25	50.39	No	S77N E79S
NN-7491	72	0.10	36.49	ND	P232N V234T
NN-7492	73	0.56	66.36	Yes (Partial) <50%	Q316N D318S
NN-7493	74	0.29	43.97	Yes (Partial) >50%	K60N V62S
NN-7494	75	0.17	35.80	Yes	S237N V239T
NN-7495	76	0.93	83.16	Yes	T297N Y299T
NN-7496	77	0.16	41.62	Yes	I217N K219S
NN-7497	78	0.00	ND	ND	T330N F332T
NN-7498	79	0.00	ND	ND	H222N I224T
NN-7499	80	0.27	49.59	Yes (Partial) >50%	S90N
NN-7500	81	0.00	ND	ND	K168N I170T
NN-7501	82	0.08	26.70	No	T66N G68S
NN-7502	83	0.31	51.81	Yes	V91N Q93T
NN-7503	84	0.24	35.44	Yes	G347N Y349T

NN-7504	85	1.44	25.45	ND	V82N F84S
NN-7505	86	0.41	63.76	Yes	Y299N N301T
NN-7506	87	0.17	33.62	Yes	Y238N
NN-7507	88	0.73	67.04	Yes	E328N
NN-7508	89	0.35	53.98	No	E342N A344S
NN-7509	90	0.13	36.36	ND	V89N V91T
NN-7510	91	0.67	60.73	Yes	R272N A274S
NN-7511	92	0.55	62.51	Yes	G311N D313S
NN-7512	93	0.26	33.88	No	D246N E248S
NN-7513	94	0.31	56.93	Yes	K255N G257S
NN-7514	95	0.27	65.33	Yes (partial) <50%	R287N
NN-7515	96	0.46	62.59	Yes (Partial) <50%	L291N
NN-7516	97	0.39	79.50	Yes	R357N V359T
NN-7517	98	2.25	40.53	ND	R206N I208T
NN-7241	99	0.39	61.46	No	WT

Supplementary Table 2 Engineered N-linked glycosylated variant with ID, Serial Number, concentration, purity, assessment of glycosylation, and mutations.

A)



B)

Binding evaluation (FIX compound binding (iFluor 647 M.F.I.)/surface displayed scFv (iFluor 488 M.F.I.))

ID	NN-7467			NN-7478			NN-7489			NN-7505		
	10 nM	100 nM	1µM	10 nM	100 nM	1µM	10 nM	100 nM	1µM	10 nM	100 nM	1µM
NN-2003-L1	0.063	0.074	0.082	0.066	0.094	0.111	0.053	0.060	0.072	0.061	0.126	0.417
NN-2003-L2	0.056	0.062	0.101	0.058	0.078	0.092	0.046	0.052	0.064	0.056	0.142	0.191

Supplementary Figure 2 Binding study of NN-2003-L1 and NN-2003-L2 against four N-linked glycosylated fIXa variants using yeast surface display. A) shows the binding efficiency against the four variants. B) Tabular data of the binding study.

Engineered N-linked glycosylated sequences of fIXa

ID	Sequence
NN-7241	MPLLLLLPLLWAGALAEDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSY ECWCPFGFEGKNCELDVTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPF PCGRVSVSQTSKLTRAETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQF PWQVVLNGKVDAFCGGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRI IPHHNYNAAINKYNHDIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFH KGRSALVLQYLRVPLVDRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGT SFLTGIISWGEECAMKGKYGIYTKVSRYVNWIKEKTKLT
NN-7420	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVNTTVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYN HDIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPL VDRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAM KGKYGIYTKVSRYVNWIKEKTKLT

NN-7421	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNNATNKYN HDIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPL VDRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAM KGKYGIYTKVSRYVNWIKEKTKLT
NN-7422	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSNTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT
NN-7423	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVNQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT
NN-7424	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIENTSHTEQKRNVIRIIPHHNYNAAINKYN HDIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPL VDRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAM KGKYGIYTKVSRYVNWIKEKTKLT
NN-7425	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GNYSIYTKVSRYVNWIKEKTKLT
NN-7426	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGNKTTVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYN HDIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPL VDRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAM KGKYGIYTKVSRYVNWIKEKTKLT
NN-7427	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VNCSIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT

NN-7428	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGNVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT
NN-7429	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPNQSPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT
NN-7430	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSANNTVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT
NN-7431	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLNLSSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT
NN-7432	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGNFSWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT
NN-7433	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNNETTEHTEQKRNVIRIIPHHNYNAAINKYN HDIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPL VDRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAM KGKYGIYTKVSRYVNWIKEKTKLT
NN-7434	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECANK SKYGIYTKVSRYVNWIKEKTKLT

NN-7435	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENNKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT
NN-7436	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGNVSAFCG GSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT
NN-7437	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLNESQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT
NN-7438	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLANNTKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT
NN-7439	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVNGSDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT
NN-7440	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKNDSFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT
NN-7441	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELNESLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT

NN-7442	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNNNSAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT
NN-7443	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDANC SGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT
NN-7444	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGNHTGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT
NN-7445	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNNSNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT
NN-7446	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKNGSFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT
NN-7447	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV NRTTCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT
NN-7448	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKNTSYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAM KGKYGIYTKVSRYVNWIKEKTKLT

NN-7449	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHNGSRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT
NN-7450	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMN GTYGIYTKVSRYVNWIKEKTKLT
NN-7451	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVNNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPL VDRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAM KGKYGIYTKVSRYVNWIKEKTKLT
NN-7452	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKNNS DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT
NN-7453	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRNFSKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT
NN-7454	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIANKTYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT
NN-7455	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHNESKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT

NN-7456	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTNGSRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT
NN-7457	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVNFSCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT
NN-7458	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTNQSRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT
NN-7459	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFNESGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT
NN-7460	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSNDSKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT
NN-7461	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPNVSFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT
NN-7462	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVNITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYN HDIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPL VDRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAM KGKYGIYTKVSRYVNWIKEKTKLT

NN-7463	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDNPTVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPL VDRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAM KGKYGIYTKVSRYVNWIKEKTKLT
NN-7464	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCQNDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT
NN-7465	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDANPSQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT
NN-7466	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGNDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT
NN-7467	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLN DTATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT
NN-7468	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGENNSEETEHTEQKRNVIRIIPHHNYNAAINKYN HDIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPL VDRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAM KGKYGIYTKVSRYVNWIKEKTKLT
NN-7469	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCNQSCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT

NN-7470	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNNASNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT
NN-7471	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNNKSIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT
NN-7472	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETENTTQKRNVIRIIPHHNYNAAINKYN HDIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPL VDRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAM KGKYGIYTKVSRYVNWIKEKTKLT
NN-7473	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNANISKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT
NN-7474	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVENGTKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYN HDIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPL VDRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAM KGKYGIYTKVSRYVNWIKEKTKLT
NN-7475	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNINETEHTEQKRNVIRIIPHHNYNAAINKYN HDIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPL VDRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAM KGKYGIYTKVSRYVNWIKEKTKLT
NN-7476	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTENESTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT

NN-7477	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSANVSQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT
NN-7478	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSNKSTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT
NN-7479	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFNSSYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT
NN-7480	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTNLT
NN-7481	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHNGTSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT
NN-7482	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETNHTEQKRNVIRIIPHHNYNAAINKYN HDIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPL VDRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAM KGKYGIYTKVSRYVNWIKEKTKLT
NN-7483	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFNKSRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT

NN-7484	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEENESTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT
NN-7485	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVENTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT
NN-7486	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVNASC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT
NN-7487	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINNYSH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT
NN-7488	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETNVTITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYN HDIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPL VDRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAM KGKYGIYTKVSRYVNWIKEKTKLT
NN-7489	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTNFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT
NN-7490	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKNCSPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT

NN-7491	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDENLTLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT
NN-7492	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCNGSSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT
NN-7493	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNNVSCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT
NN-7494	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNNYTTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPL VDRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAM KGKYGIYTKVSRYVNWIKEKTKLT
NN-7495	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFNITNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT
NN-7496	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAANNSYN HDIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPL VDRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAM KGKYGIYTKVSRYVNWIKEKTKLT
NN-7497	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGNSTLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT

NN-7498	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNN DTALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPL VDRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAM KGKYGIYTKVSRYVNWIKEKTKLT
NN-7499	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVNVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT
NN-7500	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNENWTVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYN HDIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPL VDRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAM KGKYGIYTKVSRYVNWIKEKTKLT
NN-7501	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCNESYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT
NN-7502	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSNSTTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT
NN-7503	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK NKTGIYTKVSRYVNWIKEKTKLT
NN-7504	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPANPSPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT

NN-7505	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTINNTMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT
NN-7506	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSNVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT
NN-7507	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVNGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT
NN-7508	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGENCSMK GKYGIYTKVSRYVNWIKEKTKLT
NN-7509	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRNSTSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT
NN-7510	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGNSSLVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT
NN-7511	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGNRSSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT

NN-7512	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIANKSYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT
NN-7513	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLNFSSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT
NN-7514	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DNATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT
NN-7515	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCNRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSRYVNWIKEKTKLT
NN-7516	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNAAINKYNH DIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPLV DRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMK GKYGIYTKVSNYTNWIKEKTKLT
NN-7517	EDQVDPRLIDGKGGGGSGGGGSDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELD VTCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRA ETVFPDVDYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVDAFC GGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVINITPHHNYNAAINKYN HDIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSALVLQYLRVPL VDRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAM KGKYGIYTKVSRYVNWIKEKTKLT

Supplementary Table 3 Sequences of N-linked glycosylated variants

References

1. Nooren IM, Thornton JM. Diversity of protein-protein interactions. EMBO J. 2003;22(14):3486-92.
2. Wilson IA, Stanfield RL. Antibody-antigen interactions: new structures and new conformational changes. Curr Opin Struct Biol. 1994;4(6):857-67.
3. Ostergaard H, Lund J, Greisen PJ, Kjellev S, Henriksen A, Lorenzen N, et al. A factor VIIIa-mimetic bispecific antibody, Mim8, ameliorates bleeding upon severe vascular challenge in hemophilia A mice. Blood. 2021;138(14):1258-68.
4. Kitazawa T, Igawa T, Sampei Z, Muto A, Kojima T, Soeda T, et al. A bispecific antibody to factors IXa and X restores factor VIII hemostatic activity in a hemophilia A model. Nat Med. 2012;18(10):1570-4.
5. Kohler G, Milstein C. Continuous cultures of fused cells secreting antibody of predefined specificity. Nature. 1975;256(5517):495-7.
6. Becker W, Bhattiprolu KC, Gubensak N, Zangger K. Investigating Protein-Ligand Interactions by Solution Nuclear Magnetic Resonance Spectroscopy. Chemphyschem. 2018;19(8):895-906.
7. Monaco S, Tailford LE, Juge N, Angulo J. Differential Epitope Mapping by STD NMR Spectroscopy To Reveal the Nature of Protein-Ligand Contacts. Angew Chem Int Ed Engl. 2017;56(48):15289-93.
8. Davies DR, Cohen GH. Interactions of protein antigens with antibodies. Proc Natl Acad Sci U S A. 1996;93(1):7-12.
9. Kong L, Torrents de la Pena A, Deller MC, Garces F, Sliepen K, Hua Y, et al. Complete epitopes for vaccine design derived from a crystal structure of the broadly neutralizing antibodies PGT128 and 8ANC195 in complex with an HIV-1 Env trimer. Acta Crystallogr D Biol Crystallogr. 2015;71(Pt 10):2099-108.
10. Long F, Fong RH, Austin SK, Chen Z, Klose T, Fokine A, et al. Cryo-EM structures elucidate neutralizing mechanisms of anti-chikungunya human monoclonal antibodies with therapeutic activity. Proc Natl Acad Sci U S A. 2015;112(45):13898-903.
11. Jensen PF, Larraillet V, Schlothauer T, Kettenberger H, Hilger M, Rand KD. Investigating the interaction between the neonatal Fc receptor and monoclonal antibody variants by hydrogen/deuterium exchange mass spectrometry. Mol Cell Proteomics. 2015;14(1):148-61.
12. Xu J, Cui K, Shen L, Shi J, Li L, You L, et al. Crl activates transcription by stabilizing active conformation of the master stress transcription initiation factor. Elife. 2019;8.
13. Apweiler R, Hermjakob H, Sharon N. On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. Biochim Biophys Acta. 1999;1473(1):4-8.
14. Sarkar A, Wintrode PL. Effects of glycosylation on the stability and flexibility of a metastable protein: the human serpin alpha(1)-antitrypsin. Int J Mass Spectrom. 2011;302(1-3):69-75.
15. Suga A, Nagae M, Yamaguchi Y. Analysis of protein landscapes around N-glycosylation sites from the PDB repository for understanding the structural basis of N-glycoprotein processing and maturation. Glycobiology. 2018;28(10):774-85.
16. Lombana TN, Matsumoto ML, Berkley AM, Toy E, Cook R, Gan Y, et al. High-resolution glycosylation site-engineering method identifies MICA epitope critical for shedding inhibition activity of anti-MICA antibodies. MAbs. 2019;11(1):75-93.
17. Brooks AR, Sim D, Gritzan U, Patel C, Blasko E, Feldman RI, et al. Glycoengineered factor IX variants with improved pharmacokinetics and subcutaneous efficacy. J Thromb Haemost. 2013;11(9):1699-706.
18. Rezaie AR, Esmon CT. The function of calcium in protein C activation by thrombin and the thrombin-thrombomodulin complex can be distinguished by mutational analysis of protein C derivatives. J Biol Chem. 1992;267(36):26104-9.

19. Schmidt AE, Bajaj SP. Structure-function relationships in factor IX and factor IXa. Trends Cardiovasc Med. 2003;13(1):39-45.
20. Wolberg AS, Morris DP, Stafford DW. Factor IX activation by factor XIa proceeds without release of a free intermediate. Biochemistry. 1997;36(14):4074-9.
21. Yi L, Taft JM, Li Q, Gebhard MC, Georgiou G, Iverson BL. Yeast Endoplasmic Reticulum Sequestration Screening for the Engineering of Proteases from Libraries Expressed in Yeast. Methods Mol Biol. 2015;1319:81-93.
22. Hopfner KP, Lang A, Karcher A, Sichler K, Kopetzki E, Brandstetter H, et al. Coagulation factor IXa: the relaxed conformation of Tyr99 blocks substrate binding. Structure. 1999;7(8):989-96.
23. Kabsch W, Sander C. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. Biopolymers. 1983;22(12):2577-637.
24. Fleishman SJ, Leaver-Fay A, Corn JE, Strauch EM, Khare SD, Koga N, et al. RosettaScripts: a scripting language interface to the Rosetta macromolecular modeling suite. PLoS One. 2011;6(6):e20161.
25. Leaver-Fay A, Tyka M, Lewis SM, Lange OF, Thompson J, Jacak R, et al. ROSETTA3: an object-oriented software suite for the simulation and design of macromolecules. Methods Enzymol. 2011;487:545-74.
26. Lam PV, Goldman R, Karagiannis K, Narsule T, Simonyan V, Soika V, et al. Structure-based comparative analysis and prediction of N-linked glycosylation sites in evolutionarily distant eukaryotes. Genomics Proteomics Bioinformatics. 2013;11(2):96-104.
27. Schrodinger, LLC. The PyMOL Molecular Graphics System, Version 1.8. 2015.
28. Kabsch W. Xds. Acta Crystallogr D Biol Crystallogr. 2010;66(Pt 2):125-32.
29. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. J Appl Crystallogr. 2007;40(Pt 4):658-74.
30. Liebschner D, Afonine PV, Baker ML, Bunkoczi G, Chen VB, Croll TI, et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. Acta Crystallogr D Struct Biol. 2019;75(Pt 10):861-77.
31. Johnson DJ, Langdown J, Huntington JA. Molecular basis of factor IXa recognition by heparin-activated antithrombin revealed by a 1.7-A structure of the ternary complex. Proc Natl Acad Sci U S A. 2010;107(2):645-50.
32. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. Acta Crystallogr D Biol Crystallogr. 2010;66(Pt 4):486-501.
33. Afonine PV, Grosse-Kunstleve RW, Echols N, Headd JJ, Moriarty NW, Mustyakimov M, et al. Towards automated crystallographic structure refinement with phenix.refine. Acta Crystallogr D Biol Crystallogr. 2012;68(Pt 4):352-67.
34. Zanghellini A, Jiang L, Wollacott AM, Cheng G, Meiler J, Althoff EA, et al. New algorithms and an in silico benchmark for computational enzyme design. Protein Sci. 2006;15(12):2785-94.
35. Haste Andersen P, Nielsen M, Lund O. Prediction of residues in discontinuous B-cell epitopes using protein 3D structures. Protein Sci. 2006;15(11):2558-67.
36. Buus S, Rockberg J, Forsstrom B, Nilsson P, Uhlen M, Schafer-Nielsen C. High-resolution mapping of linear antibody epitopes using ultrahigh-density peptide microarrays. Mol Cell Proteomics. 2012;11(12):1790-800.