Supplementary Material of

Structure of the zebrafish galectin-1-L2 and model of its interaction with the infectious hematopoietic necrosis virus (IHNV) envelope glycoprotein

Anita Ghosh[1,£], Aditi Banerjee[3,¥], L. Mario Amzel[2], Gerardo R. Vasta[3], and Mario A. Bianchet[1,2,*]

[1]Department of Neurology and [2]Structural Enzymology and Thermodynamics Group of the Department of Biophysics and Biophysical Chemistry, Johns Hopkins University School of Medicine, Baltimore, MD 21205.

[3]Department of Microbiology and Immunology, University of Maryland School of Medicine, Institute of Marine and Environmental Technology, Baltimore, MD 21202

[£]Current address: Department of Physiology and Biophysics, Boston University School of Medicine, 700 Albany Street, W408C, Boston, MA 02118
[¥] Current address: Department of Pediatrics, University of Maryland School of Medicine, Baltimore, MD 21201

[*]To whom correspondence should be addressed: Mario A. Bianchet, Departments of Neurology and Biophysics and Biophysical Chemistry, Johns Hopkins University School of Medicine, 725 N. Wolfe St., Baltimore, MD 21205, USA, Tel.: (410) 614-8221; E-mail: bianchet@jhmi.edu

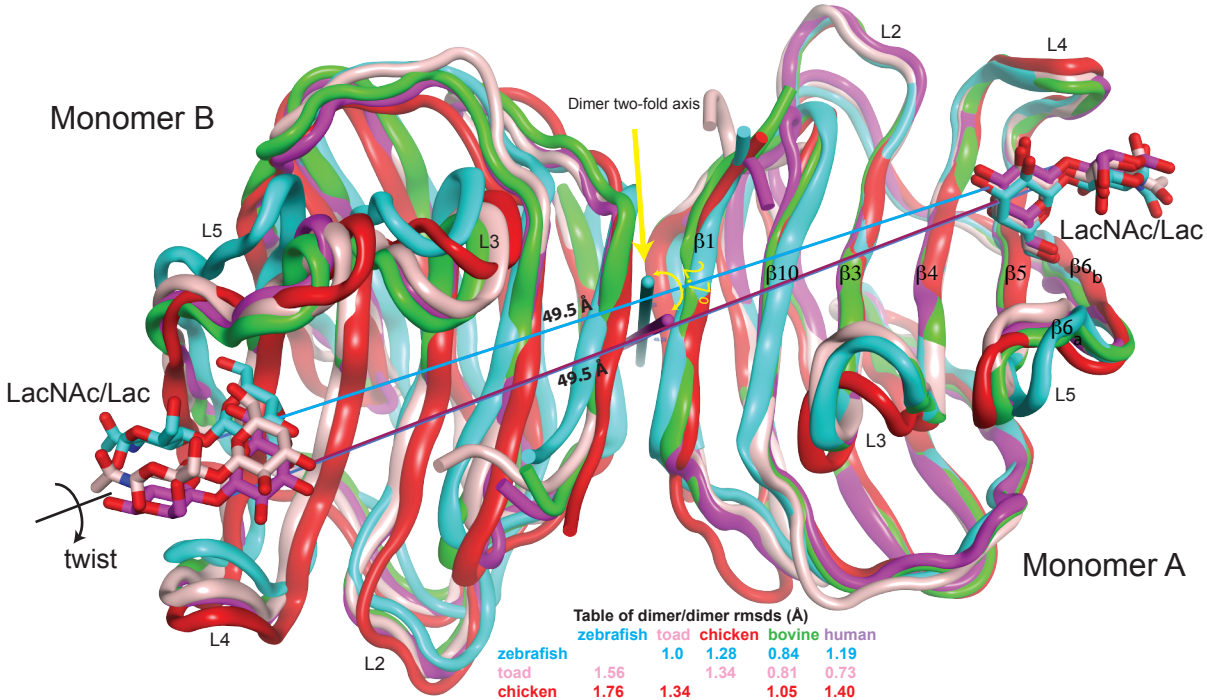
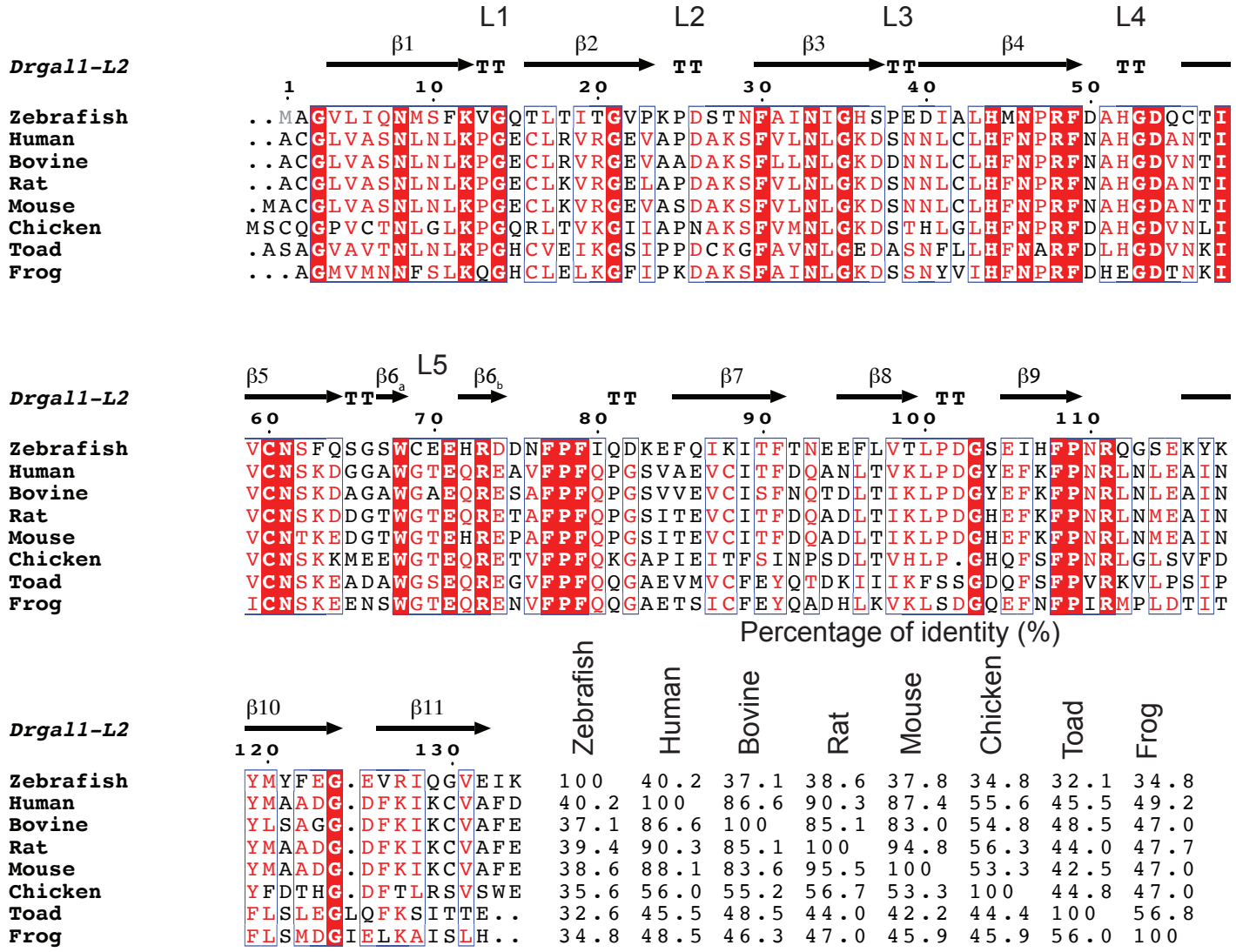


Table of dimer/dimer rmsds (Å)

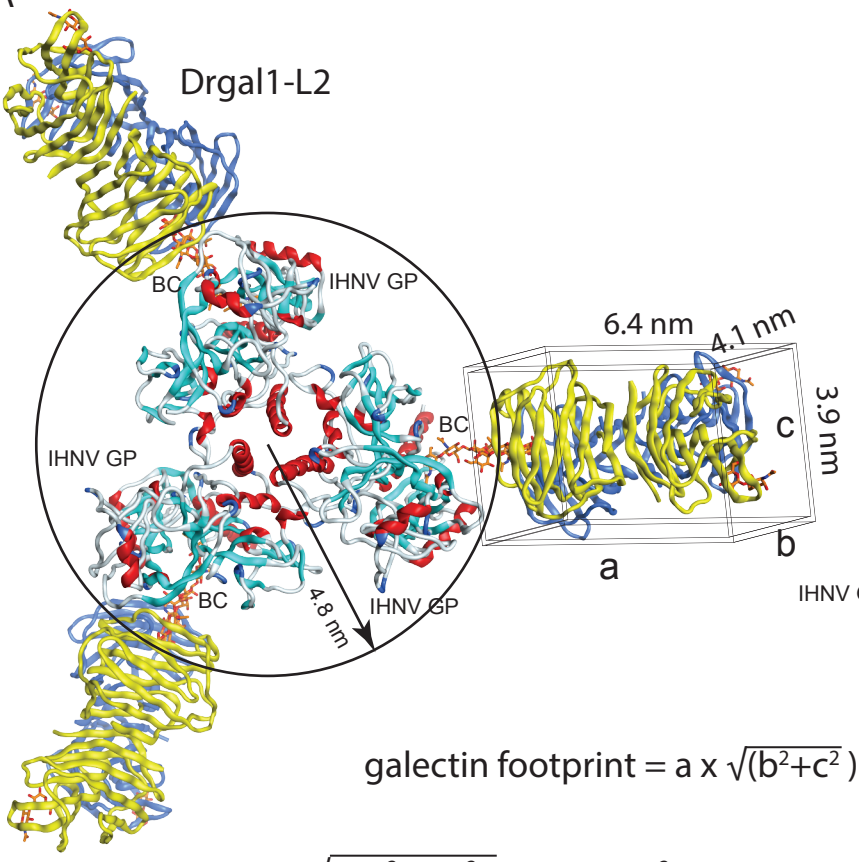
	zebrafish	toad	chicken	bovine	human
zebrafish		1.0	1.28	0.84	1.19
toad	1.56		1.34	0.81	0.73
chicken	1.76	1.34		1.05	1.40
bovine	2.42	0.96	1.56		
human	1.92	0.80	1.43	0.91	

Supplementary Figure S1. Overlay of zebrafish and known galectins-1. Toad (pink), bovine (green), and human (purple) galectins-1 dimers were aligned to Drgal1-L2 (orange), using only subunits A for the superposition. Drgal1-L2 dimer shows the smaller twist angle compared to the other galectin dimers, that cluster together. The table in the figure shows the rmsd for the pairwise alignment between dimers. The number of aligned Cα-atoms ranges from 263 to 265. The green lines connect LacNAc glycosylic-bond oxygen atoms in human and zebrafish galectins-1 of different monomers showing that the same distance between disaccharides (4.95 nm).

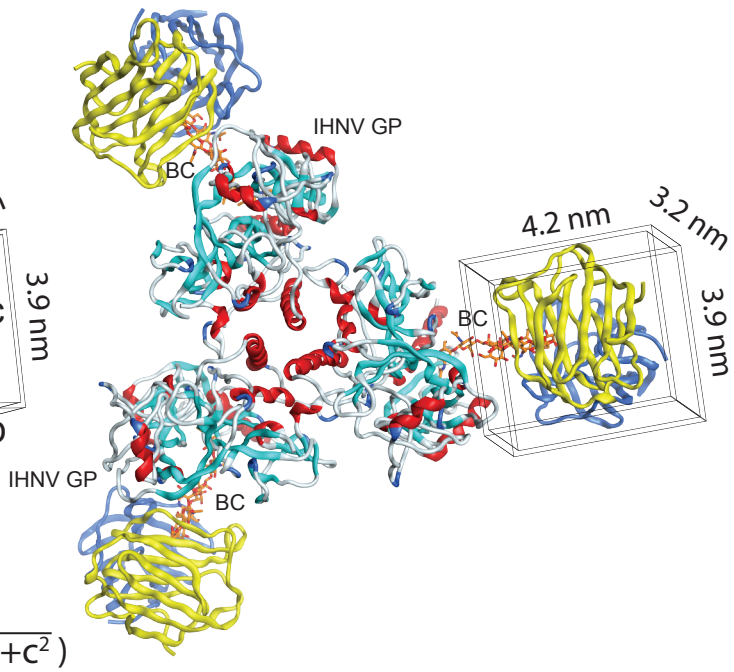

L1
L2
L3
L4
β1
β2
β3
β4
Drgal1-L2
Zebrafish
Human
Bovine
Rat
Mouse
Chicken
Toad
Frog
β5
β6a
L5
β6b
β7
β8
β9
β10
β11
Percentage of identity (%)
Zebrafish 100 40.2 37.1 38.6 37.8 34.8 32.1 34.8
Human 40.2 100 86.6 90.3 87.4 55.6 45.5 49.2
Bovine 37.1 86.6 100 85.1 83.0 54.8 48.5 47.0
Rat 39.4 90.3 85.1 100 94.8 56.3 44.0 47.7
Mouse 38.6 88.1 83.6 95.5 100 53.3 42.5 47.0
Chicken 35.6 56.0 55.2 56.7 53.3 100 44.8 47.0
Toad 32.6 45.5 48.5 44.0 42.2 44.4 100 56.8
Frog 34.8 48.5 46.3 47.0 45.9 45.9 56.0 100

Supplementary Figure S2. Sequence alignment among structurally characterized galectins-1. In the table the value at row i, column j equals to the number of residue matches between sequences i and j, divided by the length of sequence j. human galectin-1: 1GZW (Lopez-Lucendo, et al. 2004), *Rattus norvegicus* (rat) galectin-1: 3M2M (Salomonsson, et al. 2010), *Mus musculus* (mouse) galectin-1: 4LBQ (**DOI:** 10.2210/pdb4LBQ/pdb), *Gallus gallus* (chicken) galectin-1: 3DUI (López-Lucendo, et al. 2009), and *Bufo Arenarum* (toad) galectin-1: 1GAN (Bianchet, et al. 2000).

A



B



galectin footprint = $a \times \sqrt{(b^2+c^2)}$

$A_{(Drgal1)} = 6.2 \times \sqrt{(3.9^2+4.1^2)} = 35.1\ nm^2$

$A_{(1\ trimer + 3\ Drgal1)} = A_{(trimer)} + 3 \times A_{(Drgal1\text{-}L2)}$

$= 29.5 + 3 \times 35.1 = 134.5\ nm^2$

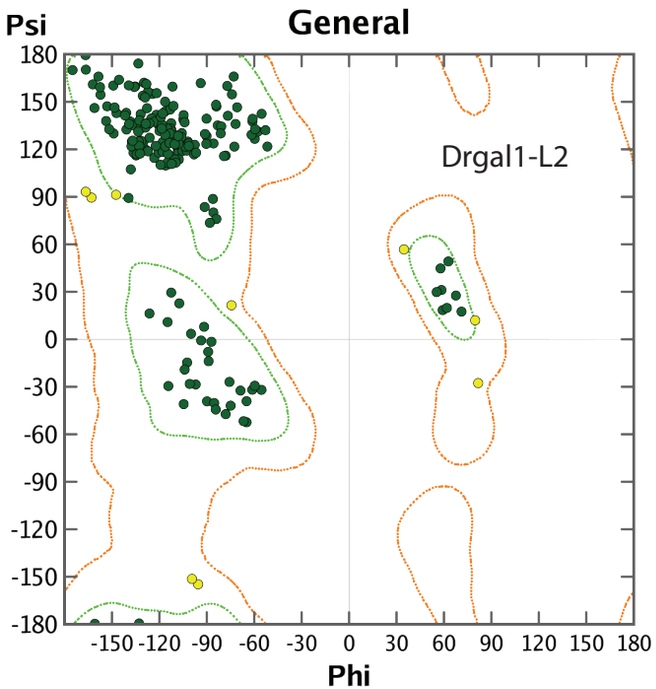
$4.2 \times \sqrt{(3.2^2+3.9^2)} = 20.5\ nm^2$

$A_{(1\ trimer + 3\ Drgal3\text{-}L1)} = A_{(trimer)} + 3 \times A_{(Drgal3\text{-}L1)}$

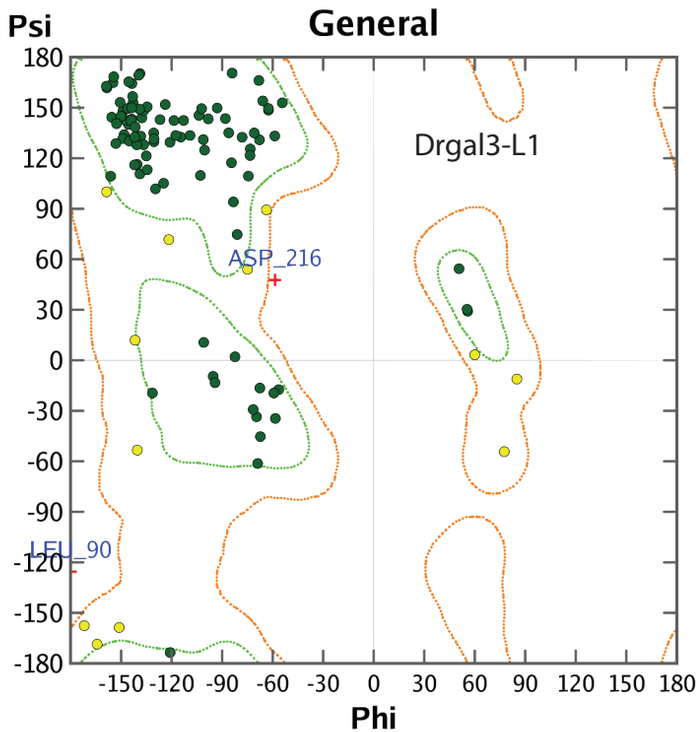
$= 29.5 + 3 \times 20.5 = 91\ nm^2$

Supplementary Figure S3. *Area calculations*. GP trimer (cyan) fully decorated by two Drgal1-L2s (yellow and blue) showing the dimensions used for the area calculations. A Drgal1-L2 dimer is contained in a 4.1 x 3.8 x 6.2 nm^3 rectangular box scribing galectin. A Drgal3-L1 in a 3.2 x 4.2 x 3.95 nm^3 rectangular box scribing the galectin. The maximum footprint area is the area of the bisecting plane of the rectangular box 35.1 nm^2 for Drgal1-L2 and 20.5 nm^2 for Drgal3-L1. A fully decorated trimer will has an are of 134 nm^2 for Drgal1-L2 and 90.1 nm^2 for Drgal3-L1. Each trimer is circumscribe by a circle of $r \approx 8.25\ nm\ \sqrt{3} = 14{,}3\ nm$.

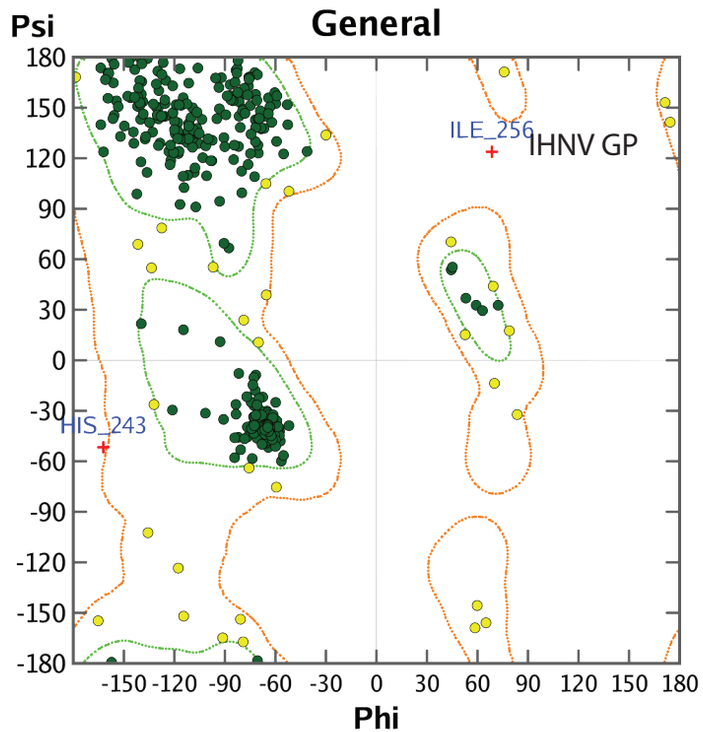
A


General
Psi
Drgal1-L2
180
150
120
90
60
30
0
-30
-60
-90
-120
-150
-180
-150 -120 -90 -60 -30 0 30 60 90 120 150 180
Phi

B


General
Psi
Drgal3-L1
ASP_216
LEU_90
180
150
120
90
60
30
0
-30
-60
-90
-120
-150
-180
-150 -120 -90 -60 -30 0 30 60 90 120 150 180
Phi

C


General
Psi
ILE_256
IHNV GP
HIS_243
180
150
120
90
60
30
0
-30
-60
-90
-120
-150
-180
-150 -120 -90 -60 -30 0 30 60 90 120 150 180
Phi

Supplementary Figure S4. *Ramachandran plots.* **(A)** Drgal1-L2, **(B)** Drgal3-L1, and **(C)** IHNV GP models

Supplementary Datos

1) Modeller python script. Save as ‘modeller.py’

++++script start at line below

```
# Modeller script for generate the IHNV GP model
from modeller import *
from modeller.automodel import *
from modeller.optimizers import conjugate_gradients, molecular_dynamics
log.verbose()    # request verbose output
env = environ()  # create a new MODELLER environment to build this model in

# directories for input atom files
env.io.atom_files_directory = './:../atom_files'
class MyLoop(loopmodel):
    def select_loop_atom(self):
        return selection(self.residue_range('375:','385:'),
                         self.residue_range('199:','193:'))
class MyModel(automodel):
    def select_atoms(self):
        return selection(self) - selection(self.residue_range('75:', '76:'))
    def special_restraints(self, aln):
        rsr = self.restraints
        at = self.atoms
        rsr.add(secondary_structure.alpha(self.residue_range('395:', '421:')))
        rsr.add(forms.gaussian(group=physical.xy_distance,
                               feature=features.distance(at['NE2:396'],at['NE2:63']),
                               mean=5.0, stdev=0.5))
        rsr.add(forms.gaussian(group=physical.xy_distance,
                               feature=features.distance(at['NE2:161'],at['NE2:396']),
                               mean=5.0, stdev=0.5))
        rsr.add(forms.gaussian(group=physical.xy_distance,
                               feature=features.distance(at['NE2:161'],at['NE2:63']),
                               mean=5.0, stdev=0.5))
    def special_patches(self,aln):
            self.patch(residue_type='DISU', residues=(self.residues['10:'],
                                                      self.residues['320:']))
```

```
a = MyModel(env,
        alnfile  = 'modelgp.ali',     # alignment filename
        knowns   = '5I2S',           # codes of the templates
        sequence = 'INHVGP')
# code of the target
a.starting_model= 4                  # index of the first modela
a.ending_model  = 4                  # index of the last model
a.make()                             # do the actual homology modelling
sele=selection(a)
a.restraints.make(sele,restraint_type='stereo',spline_on_site=False)
cg=conjugate_gradients(output='REPORT')
md=molecular_dynamics(md_return='FINAL')
for method in (cg,md,cg):
   method.optimize(sele,max_iterations=2000)
a.write(file='INHGP.pdb')
# End of python script  save in a file call model.py then run %python modeller.py
++++ cut line above
```

2) Alignment to input to Modeller. Save as 'modelgp.ali', and add the template 5I2S.pdb in the same directory in where you run modeller.py

```
++++Cut here
>P1;5I2S_X
structureX:5I2S:1 : A : 422 : A : : ::
KFTIVFPHN-QKGNWKNVPSNYHYCPSS-SDLNWHNDLIGTALQ-VKMPKSHKAIQADGW
MCHASKWVTTCDFRWYGPKYITHSIRSFTPSVEQCKESIEQTKQGTWLNPGFPPQSCG-Y
ATVTDAEAVIVQVTPHHVLVDEYTGEWVDSQFINGKCSNYICPTVHNSTT-WHSDYKVKG
LCDSNLISMDITFFSEDGELSSLGKEGTGFRSNYFAYETGGKACKMQYC-KHWGVRLPSG
VWFEMADKDLFAAARFPECPEGSSISAPSQTSVDVSLIQDVERILDYSLCQETWSKIRAG
LPISPVDLSYLAPKNPGTGPAFTIINGTLKYFETRYIRVDIAAPILSRMVGMISGTTTER
-ELWDDWAPYEDVEIGPNGVLRTSSGYKFPLYMIGHGMLDSDLHLSSKAQVFEHPHIQDA
IRKLDSFDL*
>P1;INHVGP
sequence:INHGP:38 : : 455 : : : : :
NPLFTYPEGCTLDKLSKVNASQLRCPRIFDDENRGLIAYPTSIRSLSVGNDLGEIHTQGN
HIHKVLYRTICSTGFFGGQTIEKALVEMKLSTKEAGAYDTTTAAALY----FPAPRCQWY
TDNVQNDLIFYYTTQKSVLRDPYTRDFLDSDFIGGKCTKSPCQT-HWSNVVWMGDAGIP-
ACDSSQEIKAHLFVDK------ISNRVVKATSYGHHPWGLHRACMIEFCGKQW-IRTDLG
```

DLISVEYNSGAEILSFPKCEDKTMGMRGNLDDFAY-LDDLVKASESREECLEAHAEIIST
NSVTPYLLSKFRSPHPGINDVYAMHKGSIYHGMCMTVAVDEVSKDRTTYRAHRATSFTKW
ERPFGDEWEGFHGLHGNNTTIIPDLEKYVAQYKTS--SMMMMEPMSIK-SV-PHPSILAF
YNETDLSGIS---------IRKLDSFDL*

++++Cut here

Supplementary Files

Models coordinates in a PDB format:

File 1: Model of Drgal3-L1/LacNAc complex.
Drgal3Model.pdb

File 2: Model of glycosylated IHNV GP.
IHNVGPModel.pdb

File 3: Model IHNV GP fully decorated by Drgal1-L2.
INHVGP_trimer_decorated_with_Drgal1-L2.pdb

File 4: Model IHNV GP fully decorated by Drgal3-L1.
INHVGP_trimer_decorated_with_Drgal3-L1.pdb