

Insights into the *Giardia intestinalis* Enolase and Human Plasminogen interaction

Rodrigo Aguayo-Ortiz^a, Patricia Meza-Cervantez^b, Rafael Castillo^c, Alicia Hernández-Campos^c, Laura Dominguez^{*,a}, Lilián Yépez-Mulia^{*,b}

^aFacultad de Química, Departamento de Fisicoquímica, Universidad Nacional Autónoma de México, CDMX, México 04510, México

^bUnidad de Investigación Médica en Enfermedades Infecciosas y Parasitarias-Pediatría, IMSS, México CDMX, México 06720, México

^cFacultad de Química, Departamento de Farmacia, Universidad Nacional Autónoma de México, México CDMX, México 04510, México

***Corresponding authors:**

LD: Tel.: 52(5)-622-3773. E-mail address: lauradd@unam.mx

LY-M: Tel.: 52(5)-627-6900 ext 3205, Fax: 52(5)-627-6949. E-mail address: lilianyeppez@yahoo.com

Table of contents

Contents	Page
Fig. S1	Representation of the loops involved in the <i>closed</i> and <i>open</i> state transitions of enolase.
Fig. S2	Sequence alignment of <i>Giardia intestinalis</i> enolases from different strains.
Fig. S3	Sequence alignment of enolases from different organisms.
Fig. S4	Validation plots for the model of the closed conformation of <i>Giardia intestinalis</i> enolase.
Fig. S5	Validation plots for the model of the open conformation of <i>Giardia intestinalis</i> enolase.
Fig. S6	<i>N</i> -glycosylation prediction of <i>Giardia intestinalis</i> enolase.
Fig. S7	Partial multiple sequence alignment of the proposed <i>N</i> -glycosylation site.
Fig. S8	RMSD, RMSF and Binding free energy values of GiENO-HsPLG complexes 1 and 12.
Fig. S9	Representation of the principal GiENO lysine residues involved in the interaction of HsPLG.
Fig. S10	RMSD, GiENO-HsPLG distance distribution and representative structures from GiENO mutant model simulations.

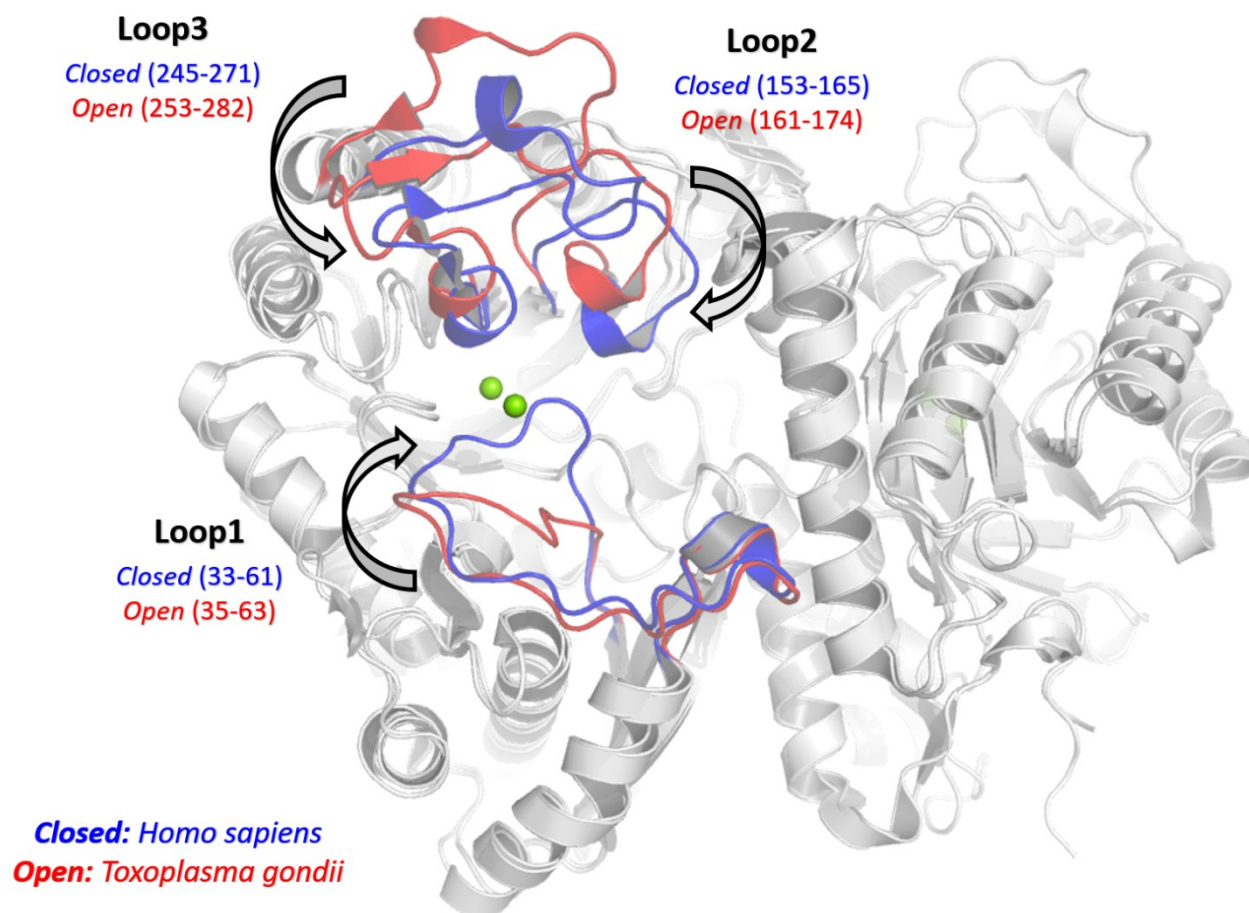


Figure S1. The superimposed structures depict the three principal connecting loops (Loop1-Loop3) associated to the transition of enolase from a *closed* (blue) to an *open* (red) state. The *Homo sapiens* (PDB ID: 3UCC)¹ and *Toxoplasma gondii* (PDB ID: 3OTR)² structures were employed to represent the *closed* and *open* conformations, respectively.

A.

	E1F9A9 (GIAIA)	C6LZY5 (GIAIB)	E2RTS6 (GIAIC)	Q8WP40 (GIAIN)	V6TRS3 (GIAIN)	V6U1Y3 (GIAIN)
E1F9A9 (GIAIA)	100					
C6LZY5 (GIAIB)	93.48	100				
E2RTS6 (GIAIC)	97.30	95.28	100			
Q8WP40 (GIAIN)	97.30	95.28	100	100		
V6TRS3 (GIAIN)	90.30	95.06	99.78 ^a	99.78 ^a	100	
V6U1Y3 (GIAIN)	93.48	100	95.28	93.48	95.06	100

^aSingle amino acid modification (Gly295Ser).

B.

E1F9A9 GIAIA	MEAPSTIKAIKARMIIDSRGTPPTEVELTTQLGVFRAACPSGASTGMHEALELRDKDPSA	60
C6LZY5 GIAIB	MEAPSTIKAIKARMIIDSRGTPPTEVELTTQLGVFRAACPSGASTGMHEALELRDKDPSA	60
E2RTS6 GIAIC	MEAPSTIKAIKARMIIDSRGTPPTEVELTTQLGVFRAACPSGASTGMHEALELRDKDPSA	60
	*****	*
E1F9A9 GIAIA	FRGKGVEQALENIRKIIAPALIGMPVRDQRAIDEKLQELDGTANKTFSILGANAVLPVSM	120
C6LZY5 GIAIB	FRGKGVEQALENIRKIIAPALIGMPVRDQRAIDEKLQELDGTAGKTF SILGANAVLPVSM	120
E2RTS6 GIAIC	FRGKGVEQALENIRKIIAPALIGMPVRDQRAIDEKLQELDGTAGKTF SILGANAVLPVSM	120
	*****	*****
E1F9A9 GIAIA	ACCLAAAREEHISLYQYLNRLANS--NNVTGARDVRLPVPCMNIIINGGKHAGNKLGPQEY	178
C6LZY5 GIAIB	ACCLAAAREEHISLYQYLNRLANNSSNTTGARDVRLPVPCMNIIINGGKHAGNKLGPQEY	180
E2RTS6 GIAIC	ACCLAAAREEHISLYQYLNRLANSNNTGARDVRLPVPCMNIIINGGKHAGNKLGPQEY	180
	*****	*****
E1F9A9 GIAIA	MIAPIKAKSFAQAMQMGCEVYSKLKEILKDEFGLAATAVGDEGGFAPNVADPEVPLKYIT	238
C6LZY5 GIAIB	MIAPTKAKNFMQAMQMGCEVYSKLKDILKEEFGLAATAVGDEGGFAPNVADPEVPLKYIT	240
E2RTS6 GIAIC	MIAPTKAKNFMQAMQMGCEVYSKLKEILKDEFGLAATAVGDEGGFAPNVADPEVPLKYIT	240
	**** *	*****
E1F9A9 GIAIA	KAIDACGYNGLMGICIDAAASEFYSETENVYDLNFKDPENPNKVSGAELANIYLDWCRRY	298
C6LZY5 GIAIB	KAIDACGYSGRMGICIDAAASEFYFKDENVYDLNFKDPENPNKVSGAELANIYLGWCKRY	300
E2RTS6 GIAIC	KAIDACGYNGLMGICIDAAASEFYSKNENVYDLNFKDPENPNKVSGAELANIYLGWCKRY	300
	*****	*****
E1F9A9 GIAIA	PIVSFEDPFNEEDHATFANFLAALRQEGLEPTQVVGDDLTVSQVSHIQKAETYSKCSALLL	358
C6LZY5 GIAIB	PIVSFEDPFDEEDHATFASFLAALKQEGLEPTQVVGDDLTVSQVSHIQKAETYSKCSALLL	360
E2RTS6 GIAIC	PIVSFEDPFDEEDHTTFANFLAALRQEGLEPTQVVGDDLTVSQVSHIQKAETYSKCSALLL	360
	*****	*****
E1F9A9 GIAIA	KINQVGTITGAIDAANLAMSYGWSVMVSHRSGETEDVFIHLATALGCGQIKSGAPSRSE	418
C6LZY5 GIAIB	KINQVGTITGAIDAANLAMSYGWSVMVSHRSGETEDVFIHLAALGCGQIKSGAPSRSE	420
E2RTS6 GIAIC	KINQVGTITGAIDAANLAMSYGWSVMVSHRSGETEDVFIHLATALGCGQIKSGAPSRSE	420
	*****	*****
E1F9A9 GIAIA	RCAKYNELLRIEDSTKIPYGFEAWK	443
C6LZY5 GIAIB	RCAKYNELLRIEDSIKVPYGFDAWK	445
E2RTS6 GIAIC	RCAKYNELLRIEDSTKIPYGFEAWK	445
	*****	*****

Figure S2. (A) Percentage sequence identity and (B) sequence alignment of *Giardia intestinalis* enolases from different strains retrieved from the UniProt server.

E2RTS6 G.intestinalis	MEAPSTIKAIKARMIIDSRGTPTEVELTTQLGVF-RAACPSGASTGMHEALELRDKDPS	59
Q97QS2 S.pneumoniae	---MSIITDVYAREVLDSRGNTLEVEVYTESGAFGRGMVPSGASTGEHEAVELRDGDKS	57
Q4DZ98 T.cruzi	---MTIQKVHGREILDSRGNTVEVEVTTTELGVF-RSAVPSGASTGIHEACELRDDDKR	55
Q9NDH8 T.bruceibrucei	---MTIQKVHGREVLDSRGNTVEVEVTTTERGVF-RSAVPSGASTGVYEACELRDGDKK	55
P09104 H.sapiens	---MSIEKIWAREILDSRGNTVEVDLYTAKGLF-RAAVPSGASTGIYEALERDGDQKQ	55
	* : .* :****.* **.: * * * *. ***** :** ***** *	
E2RTS6 G.intestinalis	AFRGKGVEQALENIRKIIAPALIGM--PVRDQRAIDEKLQELDGTAGKTF SILGANAVLP	117
Q97QS2 S.pneumoniae	RYGGLGTQKAVDNVNNIIAEAIIGY--DVRDQQAIDRAMIALDGTGNK--GKLGANAILG	113
Q4DZ98 T.cruzi	RYLGKGCLNAVKNVNDVLAPALVGK--DELQQTLDKLMRDLDTGNK--SKLGANAILG	111
Q9NDH8 T.bruceibrucei	RYVGKGCLQAVKNVNEVIGPALIGR--DELKQEELDTLMLRLDTGNK--GKLGANAILG	111
P09104 H.sapiens	RYLGKGVKLAVDHINSTIAPALISSGLSVVQEKLNDLMLLELDGTENK--SKFGANAILG	113
	: * * :*:..... :. *:.. . * : * : **** * . :****:*	
E2RTS6 G.intestinalis	VSMACCLAAAREEHISLYQYLNRLANSNSSNVTGARDVRLPVPCMNIIINGGKHAGNKLGP	177
Q97QS2 S.pneumoniae	VSIIVAAAAADYLEIPLYSYLGGFN-----TKVLPTPMNIIINGGSHSDAPIAF	162
Q4DZ98 T.cruzi	CSMAISKAAAARKGVPLRYRYLAELA-----GTKEVRLPVPCFNVIINGGKHAGNALPF	163
Q9NDH8 T.bruceibrucei	CSMAISKAAAAAGVPLRYRYLASLA-----GTKEVRLPVPCFNVIINGGKHAGNALPF	163
P09104 H.sapiens	VSLAVCKAGAAERELPLYRHIAQLA-----GNSDLILPVPAFNVINGGSHAGNKLAM	165
	: . *.* : ** : : **.* :*:****.*: :	
E2RTS6 G.intestinalis	QEYMIAPTAKNFMQAMQMGCEVYSKLKEILKDEFLAATAVGDGEGFAPNVADPEVPLK	237
Q97QS2 S.pneumoniae	QEFMILPVGAPTFKEALRYGAEIFHALKKILKSRG--LETAVGDGEGFAPRFEGTEDGVE	220
Q4DZ98 T.cruzi	QEFMIAPVKAGSFNEALRMGAEVYHSLKSIKKKKYQDAVNVGDGEGFAPPITDINEPLP	223
Q9NDH8 T.bruceibrucei	QEFMIAPVKATSFSEALRMGSEVYHSLRGIKKKKYQDAVNVGDGEGFAPPKIDINEPLP	223
P09104 H.sapiens	QEFMILPVGAESEFRDAMRLGAEVYHTLKGVIKDKYKGDATNVGDGEGFAPNILENSEALE	225
	.* * . * . * :*: *.*: : : :*. . *** . . :	
E2RTS6 G.intestinalis	YITKAIDACGYNG--LMGICIDAAASEFYSKNENVYDLNFKDPENPNKVSGAELANIYLG	295
Q97QS2 S.pneumoniae	TILAAIEAAGYVPGKDVIFGDFCASSEFYDKERKVDYDTKFEGEGAAVRTSAEQIDYLEE	280
Q4DZ98 T.cruzi	ILMEAIEQAGHKG--RFAICMDSAASETYDENKKQYNLTFKSPE-ATWVTAKQLAETYAK	280
Q9NDH8 T.bruceibrucei	ILMEAIEEAGHRG--KFAICMDCAASETYDEKKQYNLTFKSPE-PTWVTAEQLRETYCK	280
P09104 H.sapiens	LVKEAIDKAGYTE--KIVIGMDVAASEFYRDGKYDLDFKSPTDP-SRYITGDQLGALYQD	282
	: ** : .* : . * : * * * * . . . : . : :	
E2RTS6 G.intestinalis	WCRRYPIVSFEDPFDEEDHTTFANFLAALRQEGLPQVVGDDLTVSQVSHIQKAETYKSC	355
Q97QS2 S.pneumoniae	LVNKYPIITIEDGMDENDWDGWKAL TERLGK--KVQLVGDDFFVTNTDYLARGIQEGAA	337
Q4DZ98 T.cruzi	WVSEYPIVSLIEDPYDQDDFAGFAGITEALKG---KAQVVGDDLTVTNVSRIKTAIEKKAC	337
Q9NDH8 T.bruceibrucei	WAHDYPIVSIEDPYDQDDFAGFAGITEALKG---KTQIVGDDLTVTNTERIKMAIEKKAC	337
P09104 H.sapiens	FVRDYPVVSIEDPFQDDWAWSKFTANV----GIQIVGDDLTVTNPKRIERA VEKAC	337
	:::* *::* : : : *::***: *:: . : . : . :	
E2RTS6 G.intestinalis	SALLLKINQVGTITGAIDAANLAMSYGWSVMVSHRSGETEDVFIHLATALGCGQIKSGA	415
Q97QS2 S.pneumoniae	NSILIKVNIQITLTETFEAIEAKEAGYTA VVSHRSGETEDSTIADIATNAGQIKTGS	397
Q4DZ98 T.cruzi	NSLLLKINQIGTITEAIEASKFCMSNGWSVMVSHRSGETEDTYIADLVVGLGTGQIKTGA	397
Q9NDH8 T.bruceibrucei	NSLLLKINQIGTISEAIIASSKLCMENGWSVMVSHRSGETEDTYIADLVVGLGSGQIKTGA	397
P09104 H.sapiens	NCLLLKVNQIGSVTEAIQACKLAQENGWGMVSHRSGETEDTFIADLVVGLCTGQIKTGA	397
	..*::*:***:***: : : : :. * : :***** ***** **..... *****:*	
E2RTS6 G.intestinalis	PSRSERCAKYNELLRIEDST--KIPYGF- AWK-----	445
Q97QS2 S.pneumoniae	LSRTDRIAKYNQLLRIEDQLGEVAEYRGLKSFYNLKK-	434
Q4DZ98 T.cruzi	PCRGERTAKLNQLLRIEELGAHAKFGFP-AWS-----	429
Q9NDH8 T.bruceibrucei	PCRGERTAKLNQLLRIEELGAHAKFGFP-GWS-----	429
P09104 H.sapiens	PCRSERLAKYNQLMRIEELGDEARFAGH-NFRNPSVL	434
	. * : * * * * :*:***: . : :	

Figure S3. Alignment of *Giardia intestinalis*, *Streptococcus pneumoniae*, *Trypanosoma cruzi*, *Trypanosoma brucei* and *Homo sapiens* enolase sequences. The amino acids involved in the plasminogen recognition are represented in orange boxes and the loop that contains them is highlighted in the red box. The extra amino acid fragment of *G. intestinalis* is denoted with the blue mark.

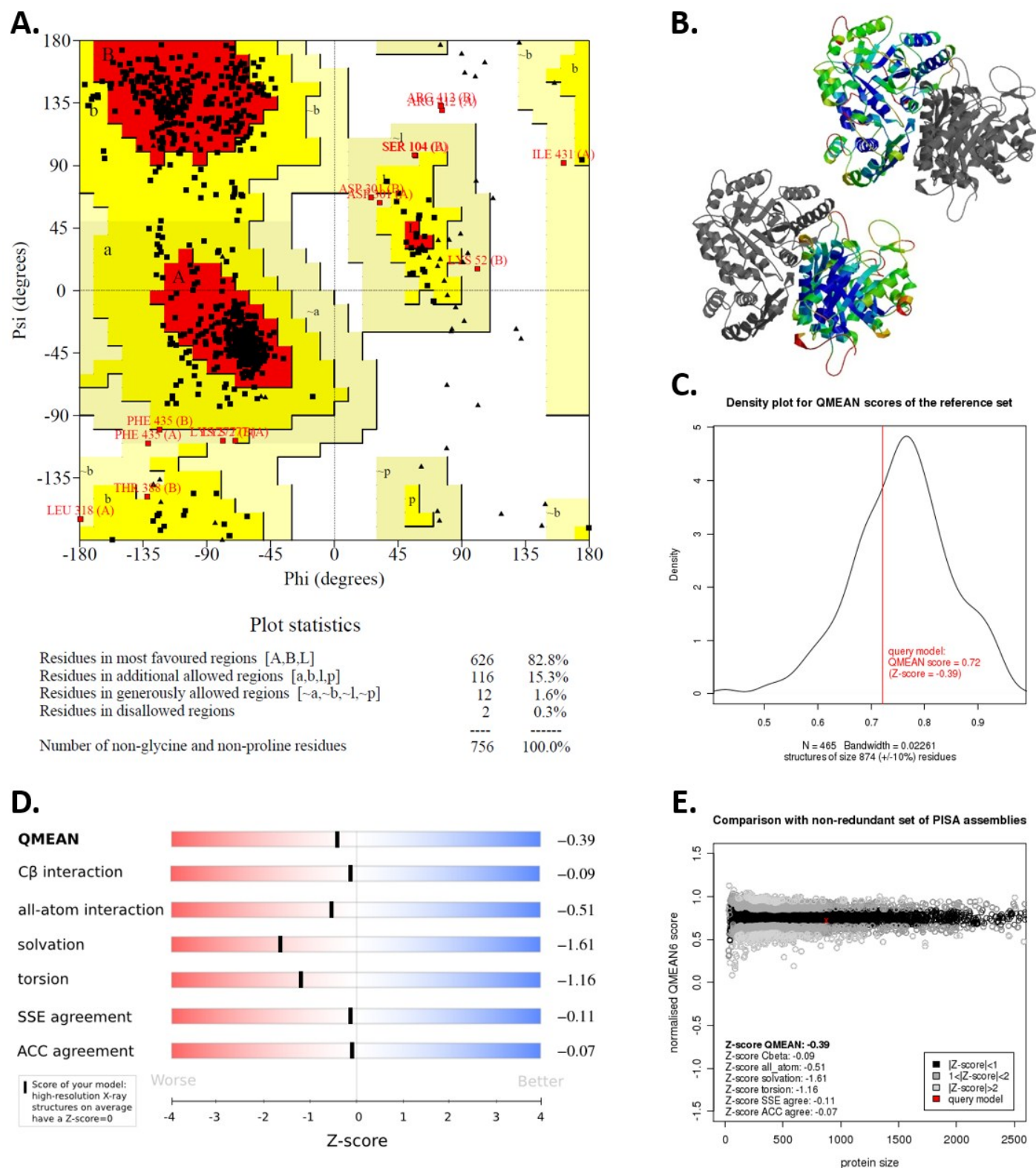


Figure S4. Validation plots for the model of the closed conformation of *Giardia intestinalis* enolase. (A) Ramachandran plot, (B) Per-residue error prediction scheme, (C) Density plot for QMEAN, (D) Z-score diagram of the individual components of QMEAN, and (E) QMEAN6 score normalized plot.

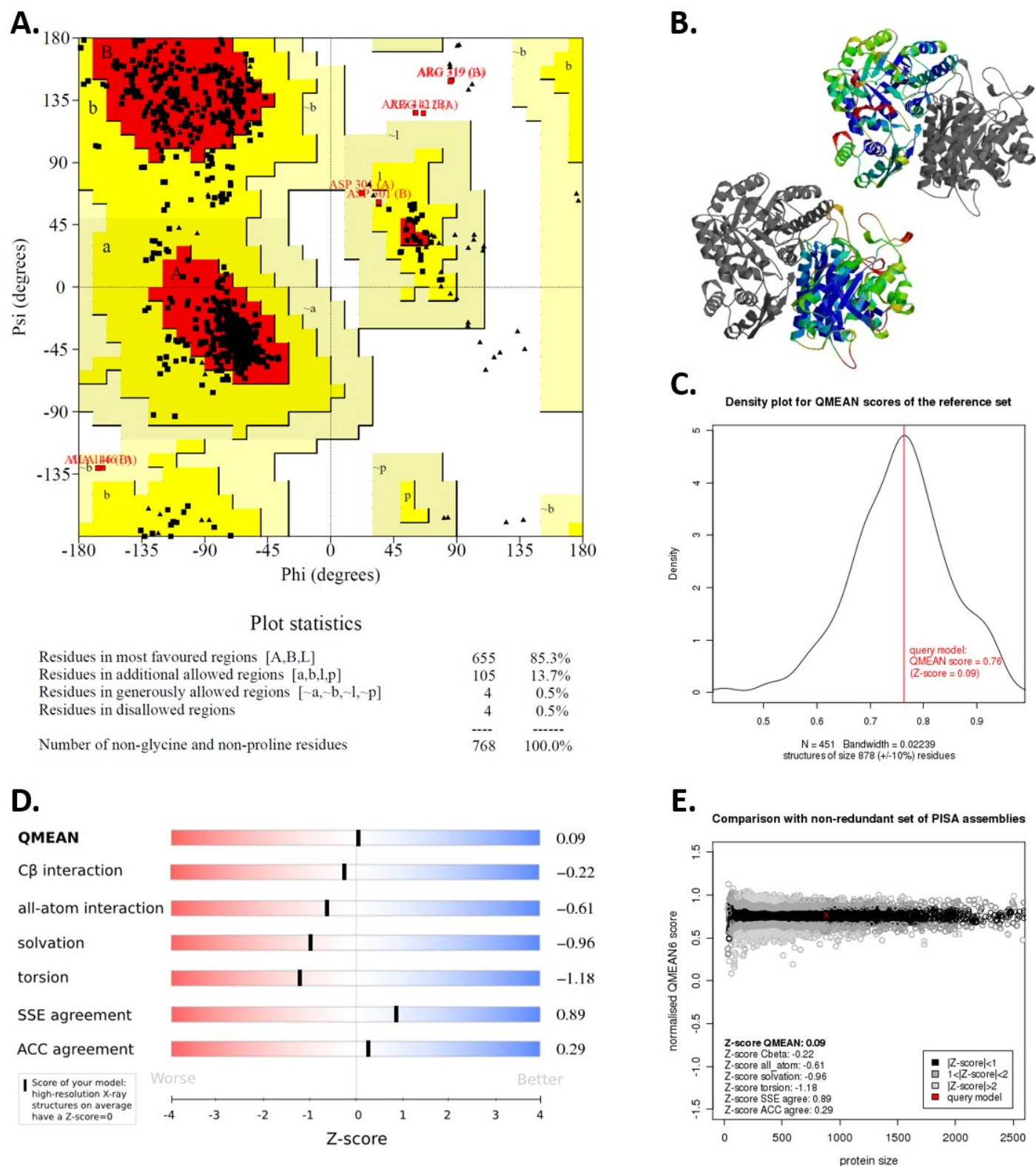


Figure S5. Validation plots for the model of the open conformation of *Giardia intestinalis* enolase. (A) Ramachandran plot, (B) Per-residue error prediction scheme, (C) Density plot for QMEAN, (D) Z-score diagram of the individual components of QMEAN, and (E) QMEAN6 score normalized plot.

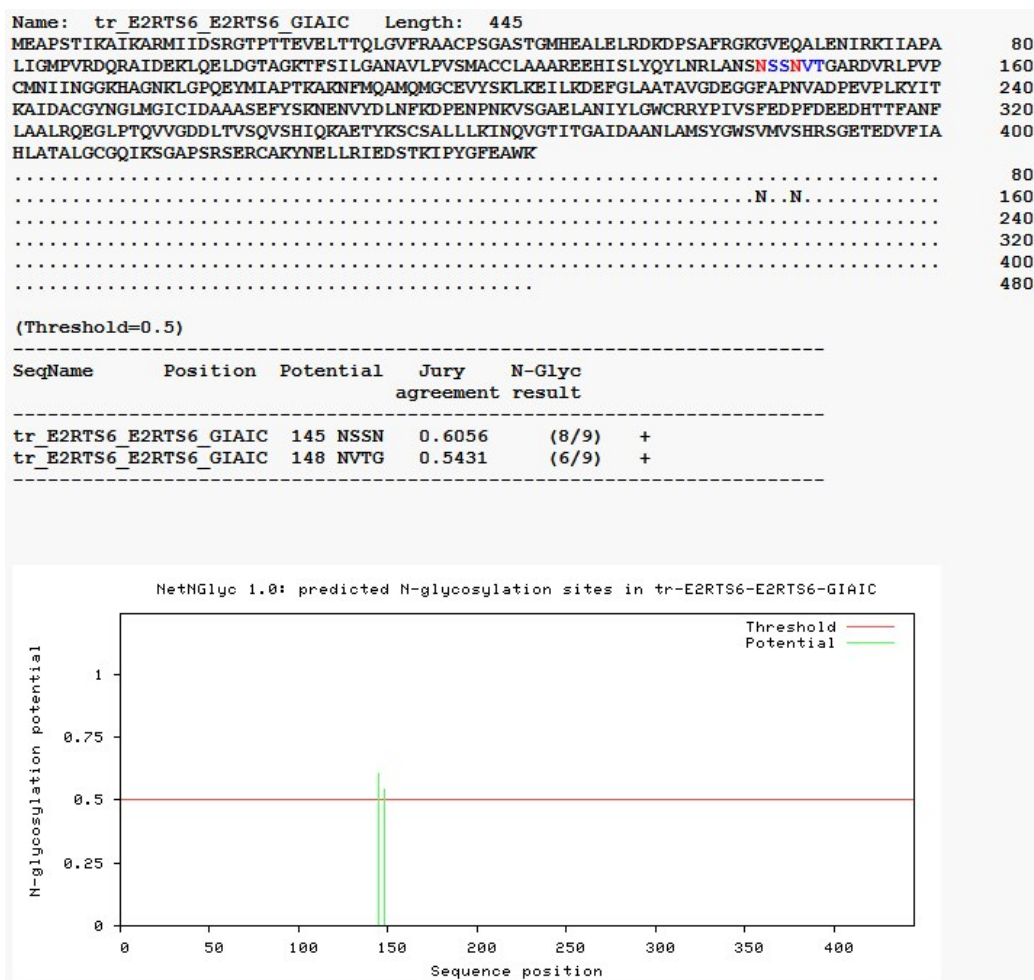


Figure S6. Prediction of *N*-glycosylation of the extra fragment loop identified in *G. intestinalis* enolase sequence.

123	CLAAAREEHISLYQYLNRANSNSSNTTGARDVRLPVP	123	C6LZY5	C6LZY5_GIAIB
123	CLAAAREEHISLYQYLNRANSNSSNVTGARDVRLPVP	123	E2RTS6	E2RTS6_GIAIC
123	CLAAAREEHISLYQYLNRANS--NNVTGARDVRLPVP	180	E1F9A9	E1F9A9_GIAIA
119	CKAGAVEKGVPLRYRIADLAGN-----SEVILPVPAFNVINGGSHAGNKLAMQEFMI	170	P06733	ENOA_HUMAN
119	CKAGAAEKGVPLRYRIADLAGN-----PDLILPVPAFNVINGGSHAGNKLAMQEFMI	170	P13929	ENOB_HUMAN
119	CKAGAAERELPLYRIHQLAGN-----SDLILPVPAFNVINGGSHAGNKLAMQEFMI	170	P09104	ENOG_HUMAN
117	CRAAAAEKGLPLYRYLAELTGH-----KEMTMPVPCFNVINGGSHAGNKLAMQEFMI	168	P51555	ENO1_ENTHI
168	CKAGAVHKGVSPLYRYIADLAGV-----DKVIMPVPAFNVINGGSHAGNKLAMQEFMI	219	E5SQX1	E5SQX1_TRISP
158	AEAGAAVQELFQYLAKAFYGGCDK--VPKKFKLPSPFFNILLGGKHAGNKLQEFMV	215	Q5EFD8	Q5EFD8_TRIVA
117	SKAAAARKGVPLRYRYLAELAGT-----KEVRLPVPCFNVINGGSHAGNKLAMQEFMI	168	K4E1A8	K4E1A8_TRYCR
118	CKAGAAEKGLPLYRYIATLAGN-----KEVIMPVPSFNVINGGSHAGNKLAMQEFMI	169	Q27655	ENO_FASHE
118	AKAGAAEKNIPLYRHFADLARN-----NEVILPVPSFNVINGGSHAGNKLAMQEFMI	169	F8UGT4	F8UGT4_MONEX
117	SKAAAARKGVPLRYRYIAALAGT-----KDRLPVPCFNVINGGSHAGNKLAMQEFMI	168	E9APW3	E9APW3_LEIMU
119	CKAGAVEKGVPLRYRIADLAGN-----AEVILPVPAFNVINGGSHAGNKLAMQEFMI	170	Q9XSJ4	ENOA_BOVIN
119	CKAGAAEKGVPLRYRIADLAGN-----PELILPVPAFNVINGGSHAGNKLAMQEFMI	170	Q32C09	ENOB_BOVIN
125	CRAGAAPNKVSLYKYLAQLAGKKS-----DQMVLPVPCFNVINGGSHAGNKLAMQEFMI	178	Q27727	ENO_PLAFA
119	ARAAAEDSGLPLYRYL---GGA-----GPMPLPVPMNVINGGEHANNSLNIQEFMI	167	B4RMD8	ENO_NEIG2
119	ARAAAEDSGLPLYRYL---GGA-----GPMPLPVPMNVINGGEHANNSLNIQEFMI	167	A9LZL4	ENO_NEIMO
119	ARAAAEDSGLPLYRYL---GGA-----GPMPLPVPMNVINGGEHANNSLNIQEFMI	167	Q9JZ53	ENO_NEIMB
119	CKAGAAEKGVPLRYRIADLAGN-----HDLVLPVPAFNVINGGSHAGNKLAMQEFMI	170	P25704	ENOB_RABIT
119	AKAAAAAGKMPLYEHIAELNGTP-----GKYSMPVPMNIIINGGEHADNNVDIQEFMI	171	P0A6P9	ENO_ECOLI
118	AKAAADSAPLPLFRYV---GGP-----NAHILPVPMNIIINGGEHADNTAVDIQEFMI	166	P96377	ENO_MYCTU
119	AKAAAAAGKMPLYEHIAELNGTP-----GKYSMPVPMNIIINGGEHADNNVDIQEFMI	171	C0PXD5	ENO_SALPC
119	AKAAAAAGKMPLYEHIAELNGTP-----GKYSMPVPMNIIINGGEHADNNVDIQEFMI	171	P64077	ENO_SALTI
119	AKAGAAEKGVPLRYAHISDLAGTKK-----PYVLPVFPQNVINGGSHAGGRILAFQEFMI	171	Q12560	ENO_ASPOR
119	AKAGAAEKGVPLRYAHISDLAGTKK-----PYVLPVFPQNVINGGSHAGGRILAFQEFMI	171	Q5B135	ENO_EMENI
119	AKAGAAEKGVPLRYAHISDLAGTKK-----PYVLPVFPQNVINGGSHAGGRILAFQEFMI	171	Q96X30	ENO_ASPFU

Figure S7. Partial multiple sequence alignment of the proposed *G. intestinalis* enolase *N*-glycosylation site.

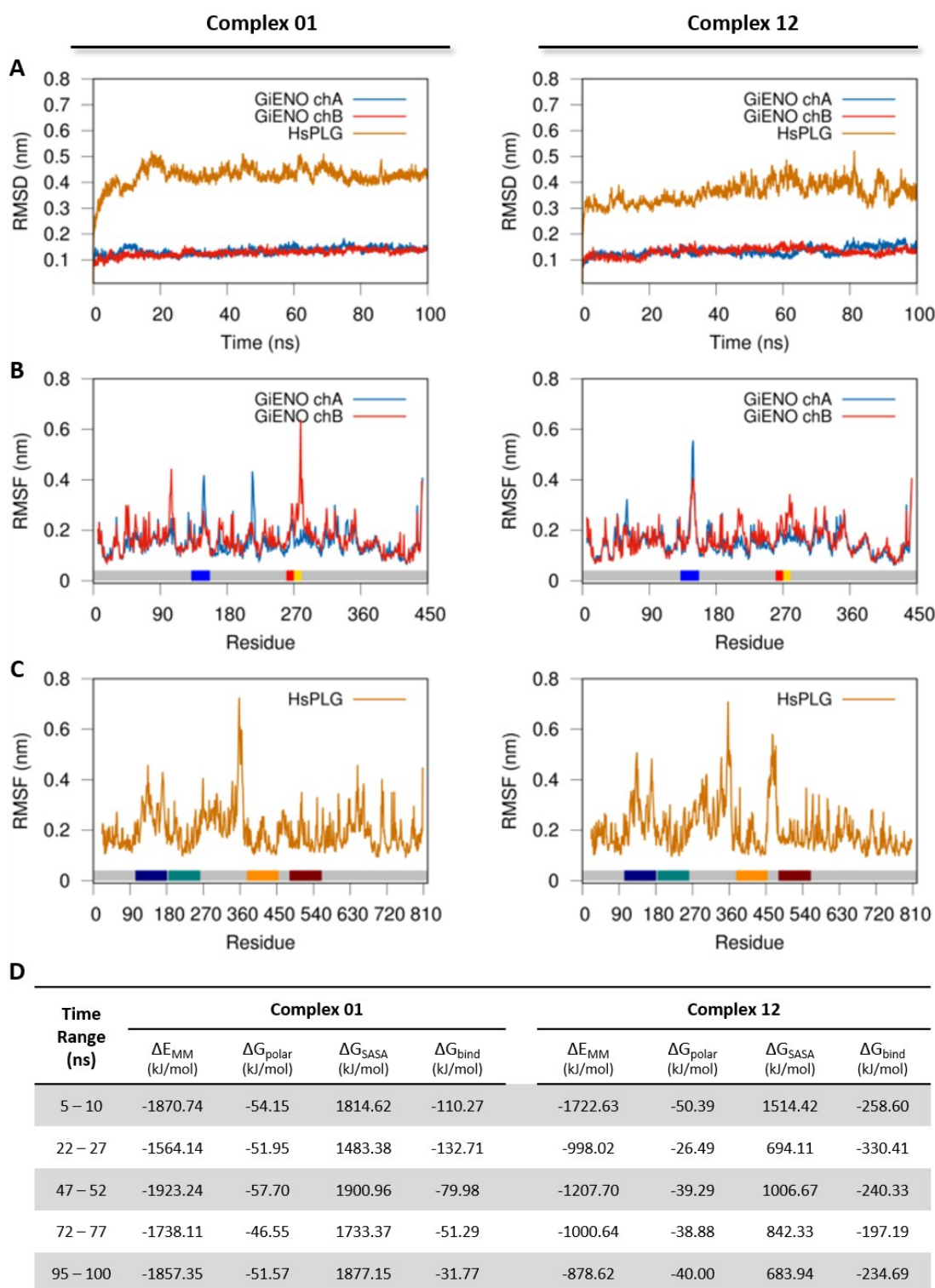


Figure S8. (A) Time evolution of the backbone-RMSD obtained from the chain A and B GiENO and HsPLG of the complexes 01 and 12. (B) Per-residue RMSF of GiENO in complexes 1 and 12, the color marks beneath the RMSF plots highlight respectively the proposed PLGBs (red), L3 loop (yellow) and N-glycosylation (blue) sites. (C) Per-residue RMSF of HsPLG in complexes 1 and 12, the color marks beneath the RMSF plots highlight respectively K1 (blue), K2 (teal), K4 (orange) and K5 (maroon) sites. (D) Binding free energy values of the GiENO-HsPLG complexes 1 and 12 calculated with MM/PBSA method.

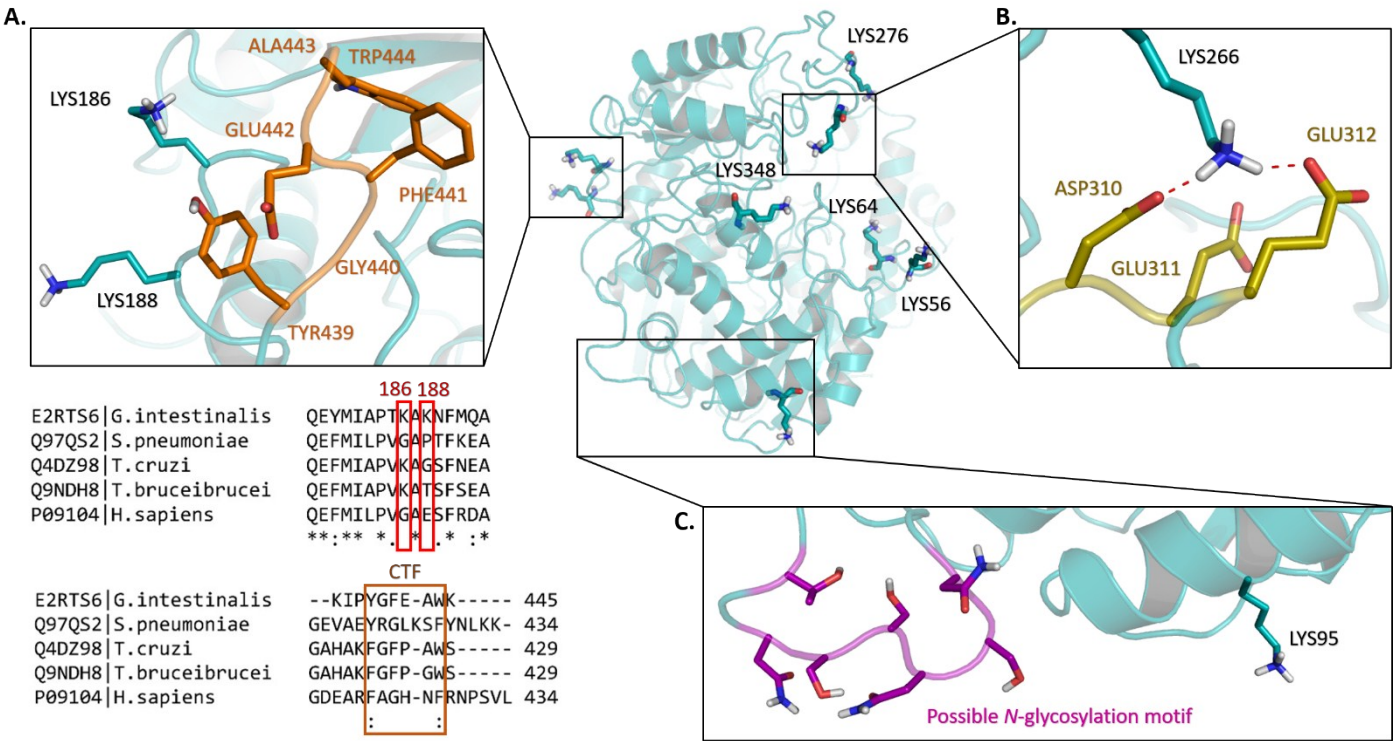


Figure S9. Representation of the principal GiENO lysine residues involved in the interaction of HsPLG. (A) Proximity of lysine residues 186 and 188 to the C-terminal fragment of GiENO. (B) Hydrogen bond interaction of Lys266 with Asp310 and Glu312. (C) Spatial representation of the distance between t Lys95 and the possible *N*-glycosylation motif.

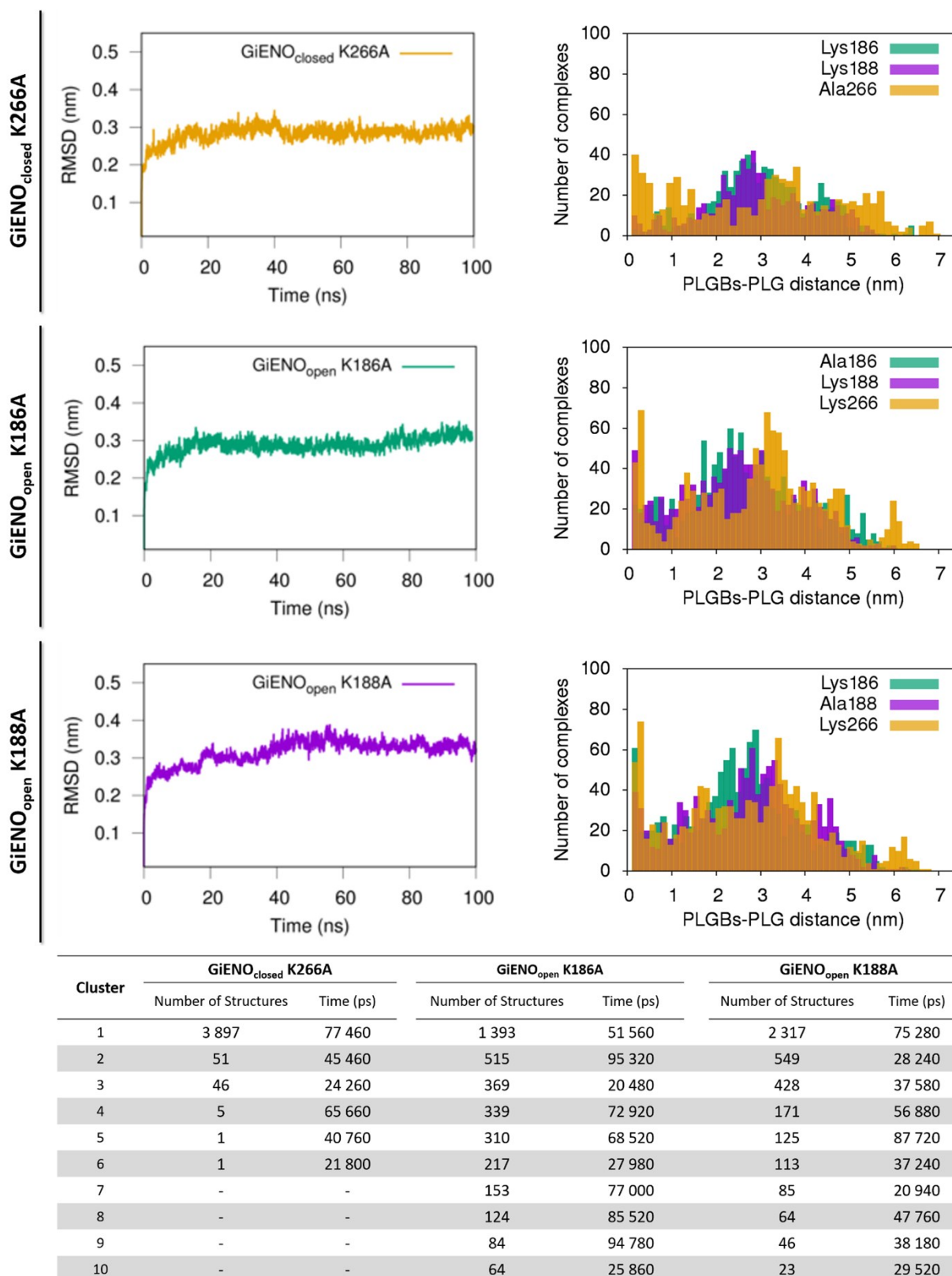


Figure S10. Time evolution of the backbone-RMSD obtained from the GiENO K266A (yellow), K186A (teal) and K188A (purple) simulated mutant models and PLGBs-HsPLG distance distribution of GiENO-HsPLG complexes (left and right plots, respectively). The table beneath the plots report the representative structures of GiENO simulations obtained from RMSD clustering.

References

1. Qin, J., Chai, G., Brewer, J. M., Lovelace, L. L. & Lebiada, L. Structures of asymmetric complexes of human neuron specific enolase with resolved substrate and product and an analogous complex with two inhibitors indicate subunit interaction and inhibitor cooperativity. *J. Inorg. Biochem.* **111**, 187–194 (2012).
2. Ruan, J. *et al.* The structure of bradyzoite-specific enolase from *Toxoplasma gondii* reveals insights into its dual cytoplasmic and nuclear functions. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **71**, 417–426 (2015).