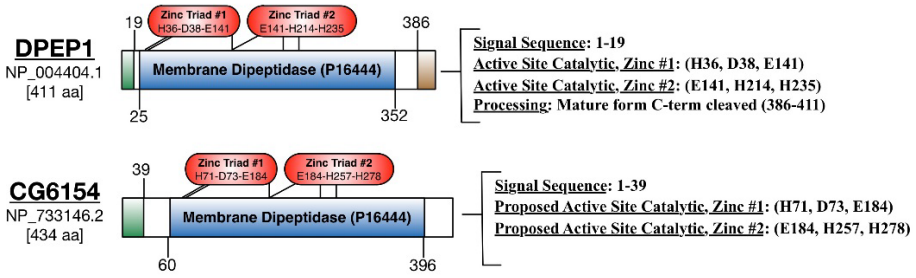
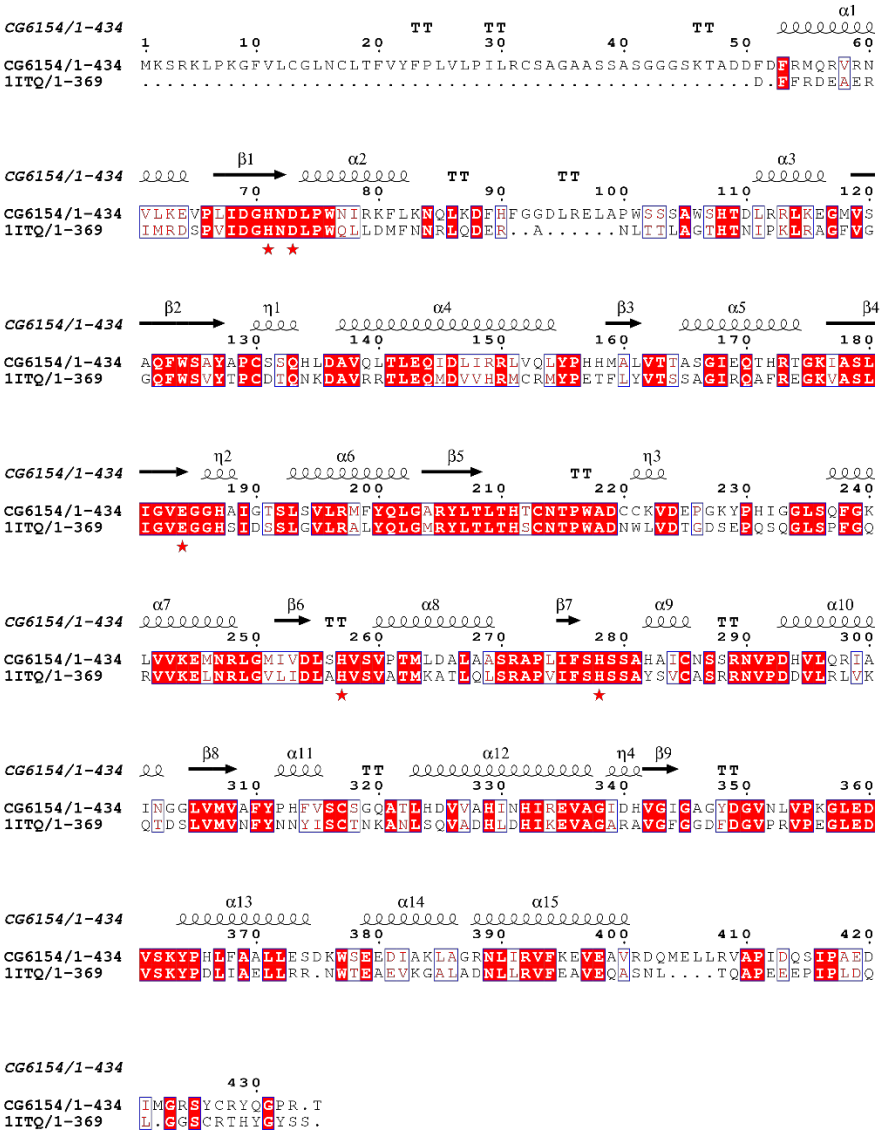
A.


DPEP1
NP_004404.1
[411 aa]
Zinc Triad #1
H36-D38-E141
Zinc Triad #2
E141-H214-H235
19
386
Membrane Dipeptidase (P16444)
25
352
Signal Sequence: 1-19
Active Site Catalytic, Zinc #1: (H36, D38, E141)
Active Site Catalytic, Zinc #2: (E141, H214, H235)
Processing: Mature form C-term cleaved (386-411)
CG6154
NP_733146.2
[434 aa]
Zinc Triad #1
H71-D73-E184
Zinc Triad #2
E184-H257-H278
39
Membrane Dipeptidase (P16444)
60
396
Signal Sequence: 1-39
Proposed Active Site Catalytic, Zinc #1: (H71, D73, E184)
Proposed Active Site Catalytic, Zinc #2: (E184, H257, H278)

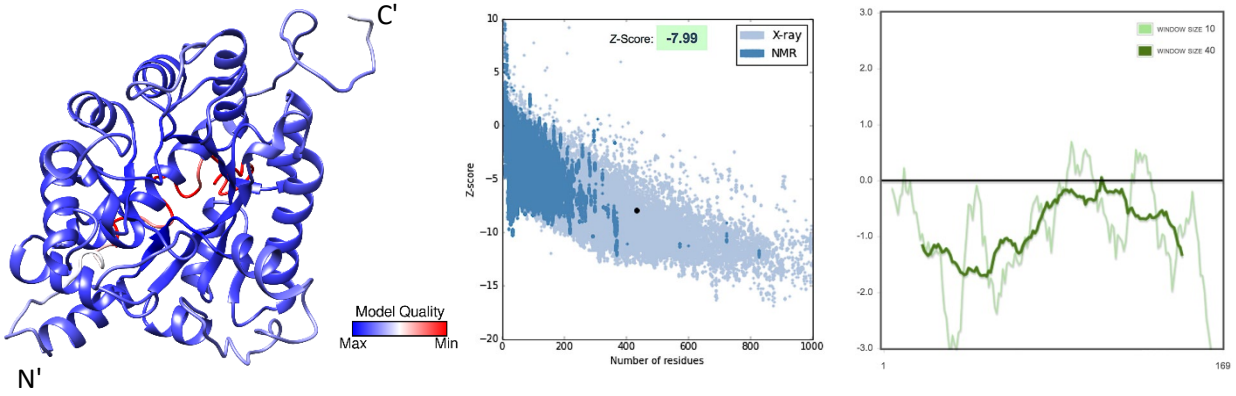
B.


CG6154/1-434
CG6154/1-434 MKSRKLPKGFVLCGLNCLTFVYFPLVLPILRCSAGAASSASGGGSKTADDFDFRMQRVRN
1ITQ/1-369 .....D.FFRDEAER
CG6154/1-434 VLKEVPLIDGHNDLPWNIRKFLKNQLKDFHFGGDLRELAPWSSSAWSHTDLRRLKEGMVS
1ITQ/1-369 IMRDSPVIDGHNDLPWQLLDMFNNRLQDFR..A.....NLTTLAGTHTNIPKLRAGFVG
CG6154/1-434 AQFWSAYAPCSSQHLDAVQLTLEQIDLIRRLVQLYPHHMALVTTASGIEQTHRTGKIASL
1ITQ/1-369 GQFWSVYTPCDTQNKDAVRRTLEQMDVVHRMCRMYPETFLYVTSSAGIRQAFREGKVASL
CG6154/1-434 IGVEGGHAIGTSLSVLRMFYQLGARYLTLTHTCNTPWADCCKVDEPGKYPHIGGLSQFGK
1ITQ/1-369 IGVEGGHSIDSSLGVLRALYQLGMRYLTLTHSCNTPWADNWLVDTGDSEPQSQGLSPFGQ
CG6154/1-434 LVVKEMNRLGMIVDLSHVSVPTMLDALAASRAPLIFSHSSAHAICNSSRNVPDHVLQRIA
1ITQ/1-369 RVVKELNRLGVLIDLAHVSVATMKATLQLSRAPVIFSHSSAYSVCASRRNVPDDVLRLVK
CG6154/1-434 INGGLVMVAFYPHFVSCSGQATLHDVVAHINHIREVAGIDHVGIGAGYDGVNLVPKGLED
1ITQ/1-369 QTDSLVMVNFYNNYISCTNKANLSQVADHLDHIKEVAGARAVGFGGDFDGVPRVPEGLED
CG6154/1-434 VSKYPHLFAALLESDKWSEEDIAKLAGRNLIRVFKEVEAVRDQMELLRVAPIDQSIPAED
1ITQ/1-369 VSKYPDLIAELLRR.NWTEAEVKGALADNLLRVFEAVEQASNL....TQAPEEEPIPLDQ
CG6154/1-434
CG6154/1-434 IMGRSYCRYQGPR.T
1ITQ/1-369 L.GGSCRTHYGYSS.

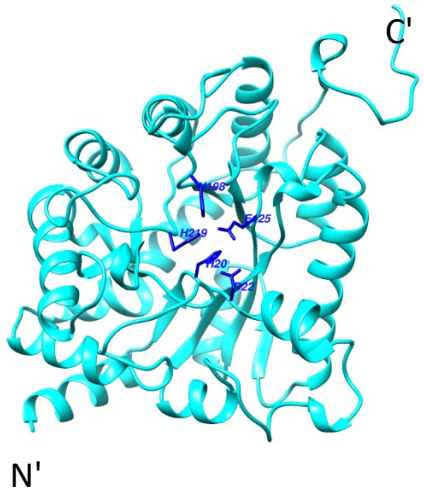
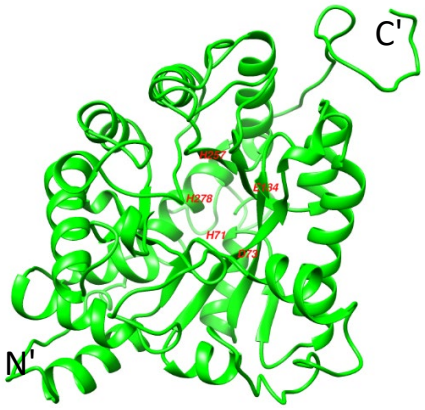
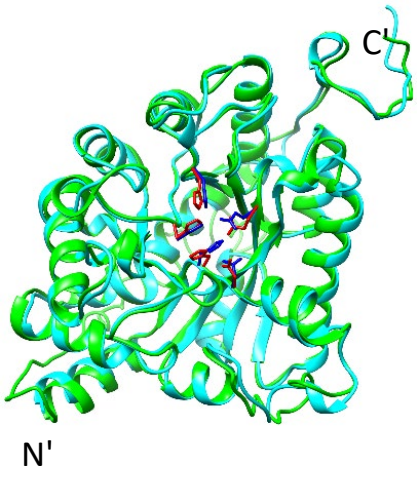
C.

```
CG6154/1-434     1 MKSRKLPKGFVLCGLNCLTFVYFPLVLPILRCSAGAASSASGGGSKTADDFDFRMQRVRNVLKEVPLIDGHN 72
1ITQ/1-369       1 ------------------------------------------------------------D-FFRDEAERIMRDSPVIDGHN 21

CG6154/1-434    73 DLPWNIRKFLKNQLKDFHFGGDLRELAPWSSSAWSHTDLRRLKEGMVSAQFWSAYAPCSSQHLDAVQLTLEQ 144
1ITQ/1-369      22 DLPWQLLDMFNNRLQDER--A------NLTTLAGTHTNIPKLRAGFVGGQFWSVYTPCDTQNKDAVRRTLEQ 85

CG6154/1-434   145 IDLIRRLVQLYPHHMALVTTASGIEQTHRTGKIASLIGVEGGHAIGTSLSVLRMFYQLGARYLTLTHTCNTP 216
1ITQ/1-369      86 MDVVHRMCRMYPETFLYVTSSAGIRQAFREGKVASLIGVEGGHSIDSSLGVLRALYQLGMRYLTLTHSCNTP 157

CG6154/1-434   217 WADCCKVDEPGKYPHIGGLSQFGKLVVKEMNRLGMIVDLSHVSVPTMLDALAASRAPLIFSHSSAHAICNSS 288
1ITQ/1-369     158 WADNWLVDTGDSEPQSQGLSPFGQRVVKELNRLGVLIDLAHVSVATMKATLQLSRAPVIFSHSSAYSVCASR 229

CG6154/1-434   289 RNVPDHVLQRIAINGGLVMVAFYPHFVSCSGQATLHDVVAHINHIREVAGIDHVGIGAGYDGVNLVPKGLED 360
1ITQ/1-369     230 RNVPDDVLRLVKQTDSLVMVNFYNNYISCTNKANLSQVADHLDHIKEVAGARAVGFGGDFDGVPRVPEGLED 301

CG6154/1-434   361 VSKYPHLFAALLESDKWSEEDIAKLAGRNLIRVFKEVEAVRDQMELLRVAPIDQSIPAEDIMGRSYCRYQGP 432
1ITQ/1-369     302 VSKYPDLIAELLRR-NWTEAEVKGALADNLLRVFEAVEQASNL----TQAPEEEPIPLDQL-GGSCRTHYGY 367

CG6154/1-434   433 R-T                                                                    434
1ITQ/1-369     368 SS-                                                                    369
```

D.



E.

DPEP1 Structure	*D. melanogaster* Model	Superimposed
		

F.	Length (AA)	Domain Architecture (Pfam, range)	Functional Residues (aligned matches in *D. melanogaster*)	Sequence ID%	Structural Overlap (RMSD)
Dipeptidase, renal (DPEP1, NP_004404.1, PDB: 1ITQ)	411	Membrane dipeptidase (PF01244) 31-349	H36, D38, E141, H214 and H235	42% ID 56% SIM	0.523 Å
DPEP (CG6154, NP_733146.2)	434	Membrane dipeptidase (PF01244) 66-393	H71, D73, E184, H257 and H278		

S6 Fig. Sequence and structural details of the modeled fly DPEP1 candidate. A. Domain architecture of DPEP1 and CG6154 and known/predicted functional residues B. Pairwise alignment of CG6154 and 1ITQ generated from structural superposition showing shared secondary structure elements and known/predicted functional residues (marked with red asterisks) C. Pairwise alignment of CG6154 and 1ITQ generated from structural superposition with conserved residues highlighted using the physiochemical color scheme (CLUSTALX) D. Validation of the CG6154 model: ProQ2 quality score mapped to a 3D model of CG6154 (left); ProSA global quality score ranking (middle) and per-residue quality graph (right) E. DPEP1 (1ITQ, cyan-blue) superimposed on the predicted structure of CG6154 (green-red) with potential matches for conserved functional residues highlighted F. Summary of features shared by DPEP1 and potential *D. melanogaster* ortholog CG6154.

.