Exercise in protein structure prediction: solution

Note. The presented slides are intended to illustrate the strategies to obtain the solution for the exercise. It is not required to submit the answers in such a format. Shorter answers that explain the obtained results were sufficient.

Using the protein structure homology modeling server SWISS-MODEL (https://swissmodel.expasy.org), predict the structures of Xrn1 enzymes of humans and yellow fever mosquitoes (accession numbers NP_061874 and EAT39382, respectively). For the sake of simplicity, for each of the proteins use only one of the templates found by SWISS-MODEL, namely the best in the default sorting according to the value of GMQE, Global Model Quality Estimation.

Which parts of the two proteins are modeled? In each of the models, identify the largest region of lower-quality prediction (coloured red in model-template alignment according to QMEAN local scores). What kind of secondary structures are predicted in these regions, α-helices, β-strands or loops? Do these regions contain insertions or deletions in modeled proteins as compared to the template?

In both cases, the same template was identified as the best one according to GMQE: PDB accession 2y35.
2y35: *Drosophila melanogaster* Xrn1

In the human Xrn1 amino acids 1-1173 were modeled (corresponding to 1-1140 of the template).
In the mosquito Xrn1 amino acids 1-1160 were modeled (same region of the template).

The largest regions of lower-quality prediction (QMEAN local scores below 0.6):
human Xrn1 : positions 356-420
mosquito Xrn1: positions 357-414.

Lower-quality prediction region in the human Xrn1 model (a loop between two α-helices):

```
Model_01  VKLSDFDREHFSEVFVDLKWFESKVGNKYLNEAAGVAAEEARNYKEKKKLKGQENSLCWTALDKNEGEMITSKDNLEDETEDDDL  425
2y35.1.A  SALTEVELDHFKEHADDLKYMNNKSEAFDM------DVGEITESQ---N-----LDSDLGALINKSML-LYDDDSEEDCSDENAV  406
```

Lower-quality prediction region in the mosquito Xrn1 model (similar loop):

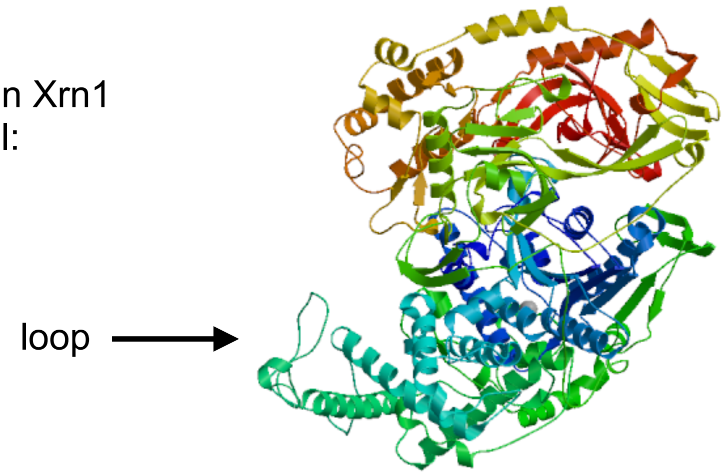
```
Model_01  RLARFDRDIFLENYTDLQYFKSKRAHNDTEAFDVIIDDIKAEAEMDCDLSALIKASEDMFSDDEDDAEPPGDIENDPEFFEKEFT  425
2y35.1.A  ALTEVELDHFKEHADDLKYMNNKSE-----AFDMDVGEITESQNLDSDLGALINKSMLLYDDDS-----EEDCSDENAVLLKEFQ  412
```

In both models, insertions as compared to the template.

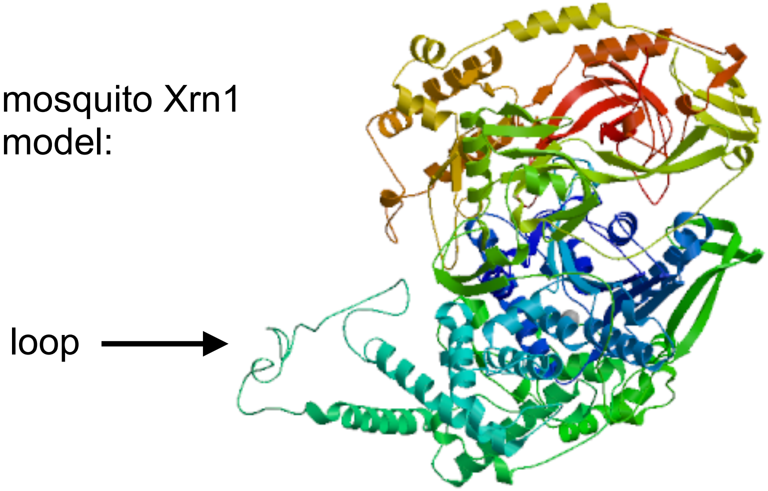
Using the protein structure homology modeling server SWISS-MODEL (https://swissmodel.expasy.org), predict the structures of Xrn1 enzymes of humans and yellow fever mosquitoes (accession numbers NP_061874 and EAT39382, respectively). For the sake of simplicity, for each of the proteins use only one of the templates found by SWISS-MODEL, namely the best in the default sorting according to the value of GMQE, Global Model Quality Estimation.

Which parts of the two proteins are modeled? In each of the models, identify the largest region of lower-quality prediction (coloured red in model-template alignment according to QMEAN local scores). What kind of secondary structures are predicted in these regions, α-helices, β-strands or loops? Do these regions contain insertions or deletions in modeled proteins as compared to the template?

human Xrn1 model:



mosquito Xrn1 model:



Lower-quality prediction region in the human Xrn1 model (a loop between two α-helices):

```
Model_01  VKLSDFDREHFSEVFVDLKWFESKVGNKYLNEAAGVAAEEARNYKEKKKLKGQENSLCWTALDKNEGEMITSKDNLEDETEDDDL  425
2y35.1.A  SALTEVELDHFKEHADDLKYMNNKSEAFDM------DVGEITESQ---N-----LDSDLGALINKSML-LYDDDSEEDCSDENAV  406
```

Lower-quality prediction region in the mosquito Xrn1 model (similar loop):

```
Model_01  RLARFDRDIFLENYTDLQYFKSKRAHNDTEAFDVIIDDIKAEAEMDCDLSALIKASEDMFSDDEDDAEPPGDIENDPEFFEKEFT  425
2y35.1.A  ALTEVELDHFKEHADDLKYMNNKSE-----AFDMDVGEITESQNLDSDLGALINKSMLLYDDDS-----EEDCSDENAVLLKEFQ  412
```

In both models, insertions as compared to the template.