|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Target Protein | Template Protein | Sequence Identity | GMQE | Q-  mean | Post Superpose RMSD | | | |
| Alpha Carbons | Back Bone | Heavy | All |
| Wild-type human neuroglobin <http://www.rcsb.org/structure/4MPM> | Horse heart ferric myoglobin <https://swissmodel.expasy.org/templates/3rj6.1> | 14.8 | 0.62 | -1.75 | 1.683 | 1.647 | 2.127 | 2.127 |
| Bar-headed goose haemoglobin <https://swissmodel.expasy.org/templates/1a4f.1> | 27.97 | 0.66 | -2.01 | 1.680 | 1.647 | 2.173 | 2.173 |
| murine neuroglobin mutant V140W <https://swissmodel.expasy.org/templates/4o2g.1> | 92.05 | 0.95 | -0.59 | 1.437 | 1.413 | 1.84 | 1.84 |
| PR domain zinc finger protein9 <https://www.rcsb.org/structure/5v3g> | Human PRDM9 allele-A ZnF Domain with Associated Recombination Hotspot <https://swissmodel.expasy.org/templates/5eh2.1> | 95.07 | 0.66 | 0.25 | 0.425 | 0.490 | 0.828 | 0.828 |
| Zinc finger and BTB domain-containing protein 17 <https://swissmodel.expasy.org/templates/2n25.1> | 43.93 | 0.09 | -0.85 | 0.61 | 0.648 | 1.005 | 1.005 |
| MouseZFP568-ZnF1-11  <https://swissmodel.expasy.org/templates/5v3m.1> | 58.27 | 0.71 | -1.34 | 0.553 | 0.570 | 0.998 | 0.998 |
| Human Cytochrome P450 1A1 <http://www.rcsb.org/structure/4I8V> | Cytochrome P450 1A2 <https://swissmodel.expasy.org/templates/2hi4.1> | 74 | 0.86 | -0.72 | 0.494 | 0.516 | 0.878 | 0.878 |
| Mammalian cytochrome P450 2B4 <https://swissmodel.expasy.org/templates/1suo.1> | 32.04 | 0.64 | -2.19 | 0.772 | 0.792 | 1.142 | 1.142 |
| Cytochrome P450 2R1 <https://swissmodel.expasy.org/templates/3czh.1> | 27.53 | 0.68 | -2.40 | 0.694 | 0.706 | 1.068 | 1.068 |

# Observations

The greater the sequence identity, the lower RMSD values from superpose. Therefore, proteins that are less homologous have higher RMSD values on superpose.

# Importance of carefully curated alignments

Curated alignments serve as templates against which an unknown amino acid sequence can be compared. The workflow enables determination of

1. *Sequence homology and its extent*
2. *Evolutionary placement of the protein sequence in question*
3. *Probable functions based on structural similarity.*
4. *Drug design needed to target specific proteins*