

BioPhysics Project

Download the FASTA sequence (UniProt ID) and convert it to PIR format.

Search for suitable templates of the sequence using BLASTp.

Select the top five templates for the next step and write the reason behind selecting these templates.

Gene Name: THC05

UniProt ID: Q13769

Download the FASTA sequence.

Fasta File:

>sp|Q13769

```
MSSESSKKRKPKVIRSDGAPAEKGKRNRSDETEGKYYSEEAEDLRDPGRDYELYKYTCQ
ELQRLMAEIQDLKSRGGKDVAIEIEERRIQSCVHFMTLKKLNRLAHIRLKKGRDQTHEAK
QKVDAYHLQLQNLLEYVMHLQKEITKCLEFKSKHEEIDLVSLLEEFYKEAPPDISKAETM
GDPHQQTARLDWELEQRKRLAEKYRECLSNKEKILKEIEVKKEYLSSLQPRLNSIMQAS
LPVQEYLFMPFDQAHKQYETARHLPPPLYVLFVQATAYGQACDKTLSVAIEGSVDEAKAL
FKPPEDSQDDESDSDAEETTKRRRPTLGVQLDDKRKEMLRHPLSVMLDLKCKDDSVL
HLTFYYLMNLNIMTVKAKVTTAMELITPISAGDLLSPDSVLSCLYPGDHGKKTNPANQY
QFDKVGILTSDYVLELGHYPYLWVQKLGGLHFPKEQPQQTVIADHSLSASHMETTMKLLK
TRVQSRLALHKQFASLEHGIVPVTSDCQYLFPKVVSRVLKVVTVAHEDYMELHFTKDIV
DAGLAGDTNLYYMALIERGTAKLQAAVVLNPGYSSIPPVFQLCLNWKGEKTNNSDDNIRA
MEGEVNVICYKELCGPWPSHQLLTNQLQRLCVLLDVYLETESHDDSVVEGPKEFPQEKMCLR
LFRGPSRMKPFKYNHPQGFFSHR
```

Converted into PIR format using:

```
from Bio import SeqIO

# Reading FASTA file
fasta_file = "input.fasta"
record = SeqIO.read(fasta_file, "fasta")

# Writing PIR file
pir_file = "output.pir"
with open(pir_file, "w") as output_handle:
    SeqIO.write(record, output_handle, "pir")
```

output.pir:

```
>P1;THC05
sequence:THC05:.....0.00: 0.00
MSSESSKKRKPKVIRSDGAPAEGRNRSDTEQEGKYYSEEAEVDLRDPGRDYELYKYTCQ
ELQRLMAEIQDLKSRGGKDVAIEIEERRIQSCVHFMTLKKLNRLAHIRLKKGRDQTAEAK
QKVDAYHLQLQNLLYEVMLHQEITKCLEFKSKHEEIDLVSLLEEFYKEAPPDISKAEVTM
GDPHQQTARLDWELEQRKRLAEKYRECLSNKEKILKEIEVKKEYLSSLQPRLNSIMQAS
LPVQEYLFMPFDQAHKQYETARHLPPPLYVLFVQATAYGQACDKTLSVAIEGSVDEAKAL
FKPPEDSQDDESDSAEEEEQTTKRRRPTLGVLDDKRKEMLRHPLSVMLDLKCKDDSVL
HLTFYYLMNLNIMTVKAKVTTAMELITPISAGDLLSPDSVLSCLYPGDHGKKTNPANQY
QFDKVGILTSDYVLELGHYPYLWVQKLGLHFPKEQPQQTVIADHLSASHMETTMKLLK
TRVQSRLALHKQFASLEHGIVPVTSDCQYLFPAKVVSRLVKWVTVAHEDYMELHFTKDIV
DAGLAGDTNLYYMALIERGTAKLQAAVVLNPGYSSIPPVFQLCLNWKGEKTNNSDDNIRA
MEGEVNVCYKELCGPWPSHQLLTNQLQRLCVLLDVYLETESHDDSVVEGPKFEPQEKMCLR
LFRGPSRMKPFKYNHPQGFFSHR*
```

Search for suitable templates of the sequence using BLASTp against the PDB database:

Similar function: Templates with known functions similar to your target protein are preferable.

Experimental structures: If available, choose templates with experimentally determined three-dimensional structures.

High sequence identity: Templates with a higher sequence identity are generally more reliable for homology modeling.

I have only two templates.

High sequence identity: Choosing templates with a high percentage of sequence identity to target protein.

Coverage: The template covers a significant portion of the target sequence.

E-value: Lower E-values indicate better homology.

Select and align chains **7ZNL_E**, and **7APK_E** to model the missing regions(these templates are selected because they have almost 100% query coverage and around 99% identity with the query).

Descriptions

Graphic Summary

Alignments

Taxonomy

Sequences producing significant alignments

Download

Select columns

Show

100

?

















☒ select all
 2 sequences selected
 [GenPept](#)
[Graphics](#)
[Distance tree of results](#)
[Multiple alignment](#)
[MSA Viewer](#)



















	Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
<input checked="" type="checkbox"/>	Chain E, THO complex subunit 5 homolog.[Homo sapiens]	Homo sapiens	1421	1421	100%	0.0	100.00%	683	7ZNL_E
<input checked="" type="checkbox"/>	Chain E, THO complex subunit 5 homolog.[Homo sapiens]	Homo sapiens	1420	1420	100%	0.0	99.71%	683	7APK_E

Using the best template structure, generate ten models. Evaluate all the models by calculating the DOPE score.

Best template structure: **7ZNL_E**

Using 7znl, generate ten models:

 7znl	18-11-2023 01:41	PDB File	7,493 KB
 Align_Model	20-11-2023 15:41	Text Document	3 KB
 Align_Model	20-11-2023 15:37	PY File	1 KB
 assess	29-10-2022 02:45	PY File	1 KB
 assess	20-11-2023 15:58	Compiled Python ...	2 KB
 automodel	29-10-2022 02:45	PY File	46 KB
 automodel	20-11-2023 15:52	Compiled Python ...	53 KB
 Multi_model	20-11-2023 16:42	Text Document	668 KB
 Multi_model	20-11-2023 16:25	PY File	1 KB
 pdb_95.pir	15-07-2022 05:56	PIR File	64,457 KB
 THCO5 -7znlE.pap	20-11-2023 15:41	PAP File	3 KB
 THCO5	20-11-2023 12:28	ALI File	1 KB
 THCO5.B99990001	20-11-2023 16:27	PDB File	426 KB
 THCO5.B99990002	20-11-2023 16:29	PDB File	426 KB
 THCO5.B99990003	20-11-2023 16:30	PDB File	426 KB
 THCO5.B99990004	20-11-2023 16:32	PDB File	426 KB

 THCO5.B99990003	20-11-2023 16:30	PDB File	426 KB
 THCO5.B99990004	20-11-2023 16:32	PDB File	426 KB
 THCO5.B99990005	20-11-2023 16:34	PDB File	426 KB
 THCO5.B99990006	20-11-2023 16:36	PDB File	426 KB
 THCO5.B99990007	20-11-2023 16:38	PDB File	426 KB
 THCO5.B99990008	20-11-2023 16:39	PDB File	426 KB
 THCO5.B99990009	20-11-2023 16:41	PDB File	426 KB
 THCO5.B99990010	20-11-2023 16:42	PDB File	426 KB
 THCO5.D00000001	20-11-2023 16:27	D00000001 File	23 KB
 THCO5.D00000002	20-11-2023 16:29	D00000002 File	19 KB
 THCO5.D00000003	20-11-2023 16:30	D00000003 File	19 KB
 THCO5.D00000004	20-11-2023 16:32	D00000004 File	22 KB
 THCO5.D00000005	20-11-2023 16:34	D00000005 File	19 KB
 THCO5.D00000006	20-11-2023 16:36	D00000006 File	21 KB
 THCO5.D00000007	20-11-2023 16:38	D00000007 File	19 KB
 THCO5.D00000008	20-11-2023 16:39	D00000008 File	17 KB
 THCO5.D00000009	20-11-2023 16:41	D00000009 File	18 KB
 THCO5.D00000010	20-11-2023 16:42	D00000010 File	23 KB

Evaluating all the models by calculating the DOPE score:

```
>> Summary of successfully produced models:
```

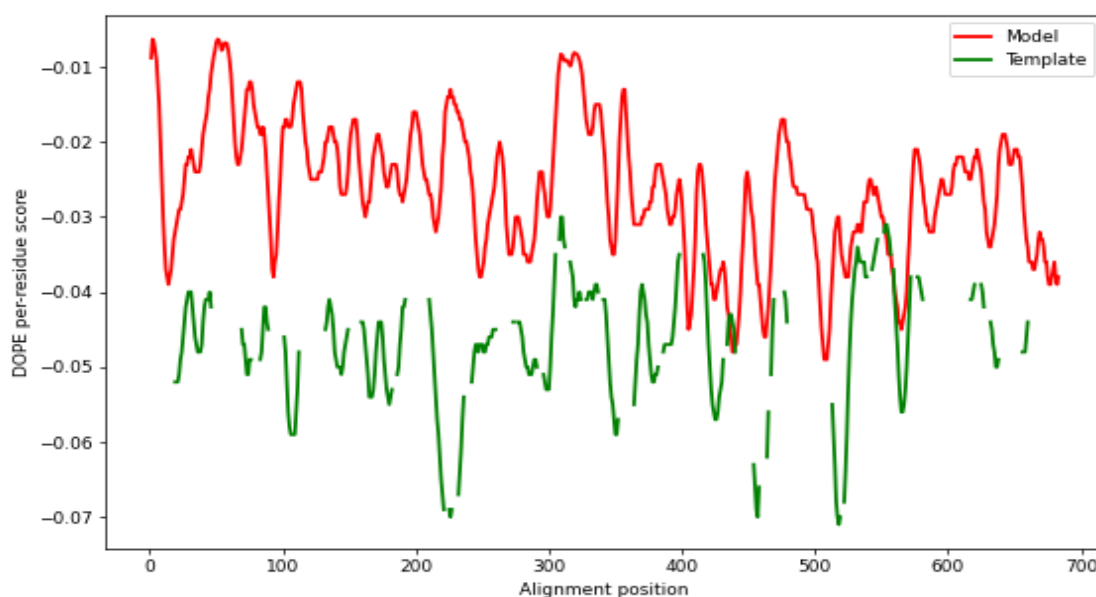
Filename	molpdf	DOPE score	GA341 score
THC05.B99990001.pdb	5361.35303	-42149.51953	0.11013
THC05.B99990002.pdb	5960.03027	-39069.76953	0.05055
THC05.B99990003.pdb	5831.55713	-42991.85547	0.08512
THC05.B99990004.pdb	5287.98389	-43196.62500	0.20044
THC05.B99990005.pdb	5772.12646	-42612.83984	0.11936
THC05.B99990006.pdb	5198.93213	-43242.58594	0.16481
THC05.B99990007.pdb	5998.61865	-42732.57422	0.06710
THC05.B99990008.pdb	6179.66113	-41441.01563	0.13697
THC05.B99990009.pdb	5930.36084	-40859.38672	0.06257
THC05.B99990010.pdb	5441.37500	-43623.36719	0.08697

Total CPU time [seconds] : 650.66

Model 10 has the best DOPE score. Because lower DOPE scores generally indicate better model quality.

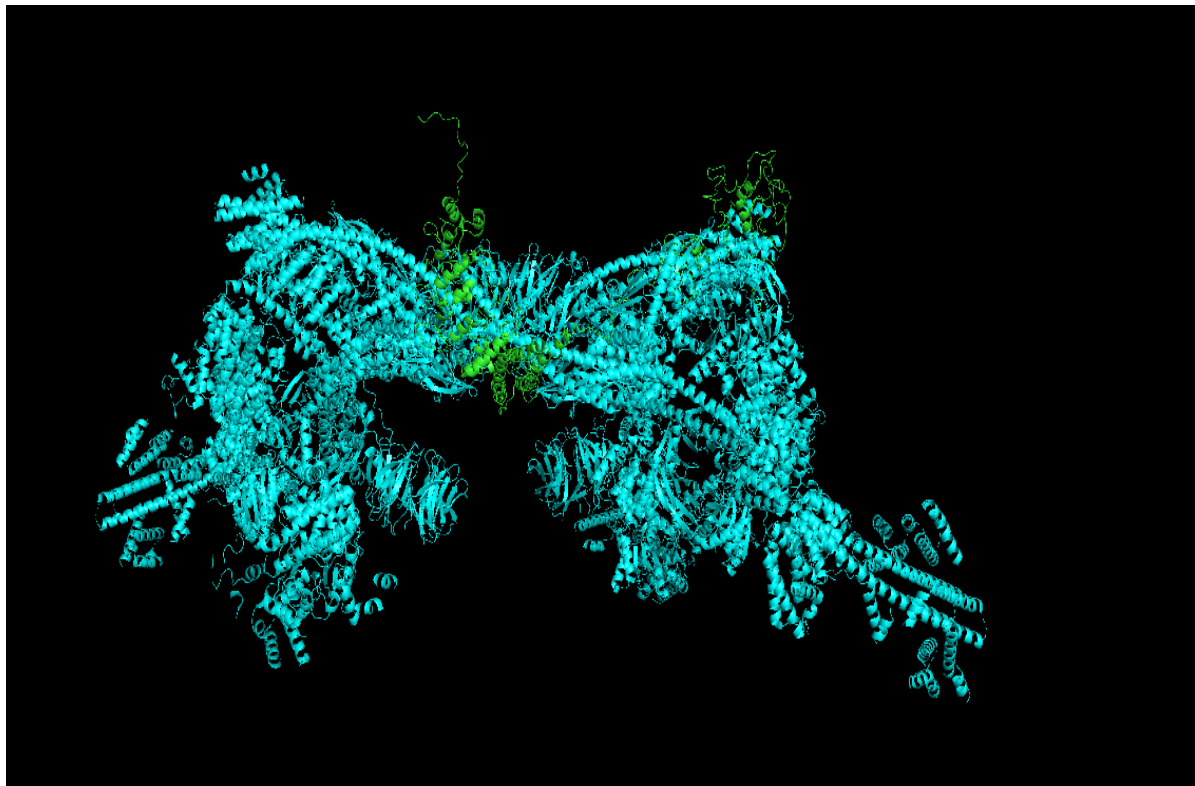
Compare the DOPE score profile of the template and best-modelled structure and attach the profile plot.

Compare the RMSD of the best model and template structure (Attach the snapshot of superimposed structures - Basic).



Comparing the RMSD of the best model and template structure (Attach the snapshot of superimposed structures - Basic)

```
CmdLoad: "" loaded as "7zn1".  
CmdLoad: "" appended into object "THC05.B99990010", state 2.  
PyMOL>align template, model  
Selector-Error: Invalid selection name "template".  
template<--  
  
PyMOL>align 7zn1 ,THC05.B99990010  
Match: read scoring matrix.  
Match: assigning 11173 x 683 pairwise scores.  
MatchAlign: aligning residues (11173 vs 683)...  
MatchAlign: score 2759.000  
ExecutiveAlign: 4145 atoms aligned.  
ExecutiveRMS: 45 atoms rejected during cycle 1 (RMSD=47.61).  
ExecutiveRMS: 18 atoms rejected during cycle 2 (RMSD=46.76).  
ExecutiveRMS: 8 atoms rejected during cycle 3 (RMSD=46.43).  
ExecutiveRMS: 1 atoms rejected during cycle 4 (RMSD=46.29).  
ExecutiveRMS: 1 atoms rejected during cycle 5 (RMSD=46.27).  
Executive: RMSD = 46.256 (4072 to 4072 atoms)
```



Refine the best-modelled structure using the Advanced Modelling technique using the top 5 templates. Generate ten models and calculate the DOPE score for each.

Got only two templates, using the advanced Modelling technique,

```
sp1;7mLE
structure7mLE_fit.pdb:381E:4538E:POL_ID 1; MOLECULE THO COMPLEX SUBUNIT 1; CHAIN A, 1, a, 1; SYNONYM THO1,NUCLEAR MATRIX PROTEIN P84,P84M5,HTREX84; ENGINEERED YES; MOL_ID 2; MOLECULE THO COMPLEX SUBUNIT 2; CHAIN B, 3, b, 1; SYNONYM THO2,HTREX120;
ENGINEERED YES; MOL_ID 3; MOLECULE THO COMPLEX SUBUNIT 3; CHAIN C, 6, c, 1; SYNONYM THO3,TEXT HOMOLOGY,HTREX45; ENGINEERED YES; MOL_ID 4; MOLECULE THO COMPLEX SUBUNIT 5 HOMOLOGY; CHAIN E, 9, e, 1; SYNONYM FUNCTIONAL SPliceOSOME-ASSOCIATED PROTEIN 79,PSAP7
NIF2/MENINGIOMA REGION PROTEIN P61.3,PLACENTAL PROTEIN 39.2, HTREX80; ENGINEERED YES; MOL_ID 5; MOLECULE THO COMPLEX SUBUNIT 6 HOMOLOGY; CHAIN F, 9, f, 1; SYNONYM FUNCTIONAL SPliceOSOME-ASSOCIATED PROTEIN 35,PSAP3 REPEAT-CONTAINING PROTEIN 58; ENGINEERED YES;
MOL_ID 6; MOLECULE THO COMPLEX SUBUNIT 7 HOMOLOGY; CHAIN G, 0, g, 0; SYNONYM FUNCTIONAL SPliceOSOME-ASSOCIATED PROTEIN 24,PSAP2 INTERACTING FACTOR 3-LIKE PROTEIN 1-BINDING PROTEIN 1,NIF3L PROTEIN 1,HTREX30; ENGINEERED YES; MOL_ID 7; MOLECULE SPliceOSOME RNA
HELICASE DDX39B; CHAIN H, 9, h, 1; SYNONYM 56 kDa UZAF65-ASSOCIATED PROTEIN,ATP-DEPENDENT RNA P47,DEAD BOX PROTEIN UAP56,HLA-B-ASSOCIATED TRANSCRIPT 1 PR EC 3.6.4.13; ENGINEERED YES;MOL_ID 1; ORGANISM_SCIENTIFIC HOMO SAPIENS; ORGANISM_COMMON HUMAN;
ORGANISM_TAXID 9606; GENE THO3, HPR1; EXPRESSION_SYSTEM TRICHOPLUSTIA NI; EXPRESSION_SYSTEM_TAXID 7111; MOL_ID 2; ORGANISM_SCIENTIFIC HOMO SAPIENS; ORGANISM_COMMON HUMAN; ORGANISM_TAXID 9606; GENE THO2, CXORF3; EXPRESSION_SYSTEM TRICHOPLUSTIA NI;
EXPRESSION_SYSTEM_TAXID 7111; MOL_ID 3; ORGANISM_SCIENTIFIC HOMO SAPIENS; ORGANISM_COMMON HUMAN; ORGANISM_TAXID 9606; GENE THO3; EXPRESSION_SYSTEM TRICHOPLUSTIA NI; EXPRESSION_SYSTEM_TAXID 7111; MOL_ID 4; ORGANISM_SCIENTIFIC HOMO SAPIENS; ORGANISM_COMMON
HUMAN; ORGANISM_TAXID 9606; GENE THO5, C22ORF9, KIA0093; EXPRESSION_SYSTEM TRICHOPLUSTIA NI; EXPRESSION_SYSTEM_TAXID 7111; MOL_ID 5; ORGANISM_SCIENTIFIC HOMO SAPIENS; ORGANISM_COMMON HUMAN; ORGANISM_TAXID 9606; GENE THO3, KIR3D5, PRCENB; EXPRESSION_SYSTEM
TRICHOPLUSTIA NI; EXPRESSION_SYSTEM_TAXID 7111; MOL_ID 6; ORGANISM_SCIENTIFIC HOMO SAPIENS; ORGANISM_COMMON HUMAN; ORGANISM_TAXID 9606; GENE THO7, NIF3LBP1; EXPRESSION_SYSTEM TRICHOPLUSTIA NI; EXPRESSION_SYSTEM_TAXID 7111; MOL_ID 7; ORGANISM_SCIENTIFIC HOMO
SAPIENS; ORGANISM_COMMON HUMAN; ORGANISM_TAXID 9606; GENE DDX39B, BAF1, UAP56; EXPRESSION_SYSTEM ESCHERICHIA COLI; EXPRESSION_SYSTEM_TAXID 562; 3.451-1.00
-----SEAVOLR-----VYTCQELQRLHAEIQDKS-----
-----VAIEEERIQSCVWFHTLKKLNLHLERLKGRODTHAQQVDAYHLQQLLYVHLNQLK-----
-----E-----RLNLEEQRLAEKRECSNEXCLKEVEKKEY
LSLSQLRLHNSIQASLPVQY-----LDQKHQVETARIHPPVYVLPQATAYGQCKTLVATGEGSDGAKAL
-----D-----DKRKLKRPILSVPLDLCKDDSDVHLTFYVYHLNIDTV
KAVVTAMELLTIPISAGDLSPDSVSLCYPDGMOKTPAPKQVQFKE-----VLSQVLELGPVLMQKLGGL
HF-----PIADHLSASHMETTHKLTKTQGRALAHQFASLEHNLVPTSDQVLPFAKVVSRVLQAMTV
AHEDVNLHFTKDIDVAGLADGTHLYMALLEETAKLQAQVLMVPSSTPPVQLCLMKGKTNNDONDIRA
REDEWVWYKELCGMPSHQLTLTQQLQVLLDVLYLETSH-----KEPQDKLRLRFGPSRNPFFKYNH
PQGFSSH*

sp1;7mPE
structure7mPE_fit.pdb:41E:523E:POL_ID 1; MOLECULE THO COMPLEX SUBUNIT 1; CHAIN A, 1, a, 1; SYNONYM THO1,NUCLEAR MATRIX PROTEIN P84,P84M5,HTREX84; ENGINEERED YES; MOL_ID 2; MOLECULE THO COMPLEX SUBUNIT 2; CHAIN B, 3, b, 1; SYNONYM THO2,HTREX120;
ENGINEERED YES; MOL_ID 3; MOLECULE THO COMPLEX SUBUNIT 3; CHAIN C, 6, c, 1; SYNONYM THO3,TEXT HOMOLOGY,HTREX45; ENGINEERED YES; MOL_ID 4; MOLECULE THO COMPLEX SUBUNIT 5 HOMOLOGY; CHAIN E, 9, e, 1; SYNONYM FUNCTIONAL SPliceOSOME-ASSOCIATED PROTEIN 79,PSAP7
NIF2/MENINGIOMA REGION PROTEIN P61.3,PLACENTAL PROTEIN 39.2, HTREX80; ENGINEERED YES; MOL_ID 5; MOLECULE THO COMPLEX SUBUNIT 6 HOMOLOGY; CHAIN F, 9, f, 1; SYNONYM FUNCTIONAL SPliceOSOME-ASSOCIATED PROTEIN 35,PSAP3 REPEAT-CONTAINING PROTEIN 58; ENGINEERED YES;
MOL_ID 6; MOLECULE THO COMPLEX SUBUNIT 7 HOMOLOGY; CHAIN G, 0, g, 0; SYNONYM FUNCTIONAL SPliceOSOME-ASSOCIATED PROTEIN 24,PSAP2 INTERACTING FACTOR 3-LIKE PROTEIN 1-BINDING PROTEIN 1,NIF3L PROTEIN 1,HTREX30; ENGINEERED YES; MOL_ID 7; MOLECULE SPliceOSOME RNA
HELICASE DDX39B; CHAIN H, 9, h, 1; SYNONYM 56 kDa UZAF65-ASSOCIATED PROTEIN,ATP-DEPENDENT RNA P47,DEAD BOX PROTEIN UAP56,HLA-B-ASSOCIATED TRANSCRIPT 1 PR EC 3.6.4.13; ENGINEERED YES; MOL_ID 8; MOLECULE THO2 ANCHOR (PUTATIVE); CHAIN X, X; ENGINEERED YES;MOL_ID
1; ORGANISM_SCIENTIFIC HOMO SAPIENS; ORGANISM_COMMON HUMAN; ORGANISM_TAXID 9606; GENE THO3, HPR1; EXPRESSION_SYSTEM TRICHOPLUSTIA NI; EXPRESSION_SYSTEM_TAXID 7111; EXPRESSION_SYSTEM_STRAIN HIS; MOL_ID 2; ORGANISM_SCIENTIFIC HOMO SAPIENS; ORGANISM_COMMON
HUMAN; ORGANISM_TAXID 9606; GENE THO2, CXORF3; EXPRESSION_SYSTEM TRICHOPLUSTIA NI; EXPRESSION_SYSTEM_TAXID 7111; EXPRESSION_SYSTEM_STRAIN HIS; MOL_ID 3; ORGANISM_SCIENTIFIC HOMO SAPIENS; ORGANISM_COMMON HUMAN; ORGANISM_TAXID 9606; GENE THO3;
EXPRESSION_SYSTEM TRICHOPLUSTIA NI; EXPRESSION_SYSTEM_TAXID 7111; EXPRESSION_SYSTEM_STRAIN HIS; MOL_ID 4; ORGANISM_SCIENTIFIC HOMO SAPIENS; ORGANISM_COMMON HUMAN; ORGANISM_TAXID 9606; GENE THO5, C22ORF9, KIA0093; EXPRESSION_SYSTEM TRICHOPLUSTIA NI;
EXPRESSION_SYSTEM_TAXID 7111; EXPRESSION_SYSTEM_STRAIN HIS; MOL_ID 5; ORGANISM_SCIENTIFIC HOMO SAPIENS; ORGANISM_COMMON HUMAN; ORGANISM_TAXID 9606; GENE THO7, NIF3LBP1; EXPRESSION_SYSTEM TRICHOPLUSTIA NI; EXPRESSION_SYSTEM_TAXID 7111; EXPRESSION_SYSTEM_STRAIN
HIS; MOL_ID 6; ORGANISM_SCIENTIFIC HOMO SAPIENS; ORGANISM_COMMON HUMAN; ORGANISM_TAXID 9606; GENE THO3, NIF3LBP1; EXPRESSION_SYSTEM TRICHOPLUSTIA NI; EXPRESSION_SYSTEM_TAXID 7111; EXPRESSION_SYSTEM_STRAIN HIS; MOL_ID 7;
ORGANISM_SCIENTIFIC HOMO SAPIENS; ORGANISM_COMMON HUMAN; ORGANISM_TAXID 9606; GENE DDX39B, BAF1, UAP56; EXPRESSION_SYSTEM ESCHERICHIA COLI; EXPRESSION_SYSTEM_TAXID 562; EXPRESSION_SYSTEM_STRAIN BL21 DE3 RIL; MOL_ID 8; ORGANISM_SCIENTIFIC HOMO SAPIENS;
ORGANISM_TAXID 9606; EXPRESSION_SYSTEM TRICHOPLUSTIA NI; EXPRESSION_SYSTEM_TAXID 7111; EXPRESSION_SYSTEM_STRAIN HIS; 3.381-1.00
-----SEVOLD-----VYTCQELQRLHAEIQDKS-----
-----VAIEEERIQSCVWFHTLKKLNLHLERLKGRODTHAQQVDAYHLQQLLYVHLNQLK-----
-----E-----RLNLEEQRLAEKRECSNEXCLKEVEKKEY
LSLSQLRLHNSIQASLPVQY-----LDQKHQVETARIHPPVYVLPQATAYGQCKTLVATGEGSDGAKAL
-----D-----DKRKLKRPILSVPLDLCKDDSDVHLTFYVYHLNIDTV
KAVVTAMELLTIPISAGDLSPDSVSLCYPDGMOKTPAPKQVQFKE-----VLSQVLELGPVLMQKLGGL
HF-----PIADHLSASHMETTHKLTKTQGRALAHQFASLEHNLVPTSDQVLPFAKVVSRVLQAMTV
AHEDVNLHFTKDIDVAGLADGTHLYMALLEETAKLQAQVLMVPSSTPPVQLCLMKGKTNNDONDIRA
REDEWVWYKELCGMPSHQLTLTQQLQVLLDVLYLETSH-----KEPQDKLRLRFGPSRNPFFKYNH
PQGFSSH*

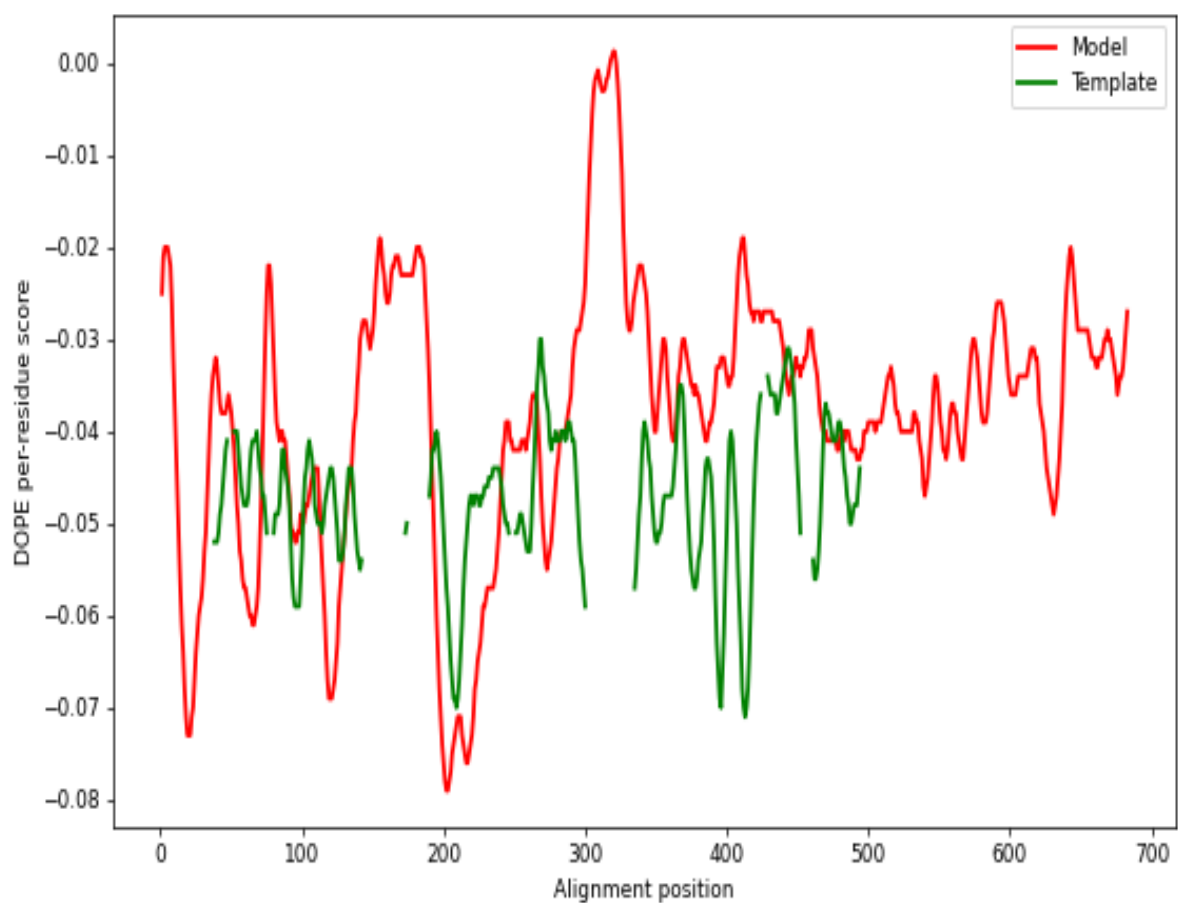
sp1;THO5
sequence:THO5: 1 11 0.00 0.00
VSESSKYNPHTKIDVAGLADGTHLYMALLEETAKLQAQVLMVPSSTPPVQYSEAVOLRDRDVEVYKTCQELQRLHAEIQDKSR
GKQDAIEEERIQSCVWFHTLKKLNLHLERLKGRODTHAQQVDAYHLQQLLYVHLNQLKLEF
KSKHEDLVSLFEKAPPDTSKAEVNDHPQGLARLDNELFQGRALAEKRECSNEXCLKEVEKKEY
LSLSQLRLHNSIQASLPVQYFPPFQKHQVETARIHPPVYVLPQATAYGQCKTLVATGEGSDGAKAL
PKPPPSQSDSDSDAEIEQTCKRRPTLQVLDKKRKLKRPILSVPLDLCKDDSDVHLTFYVYHLNIDTV
KAVVTAMELLTIPISAGDLSPDSVSLCYPDGMOKTPAPKQVQFKEVLTLSQVLELGPVLMQKLGGL
WPKREDPQVTSKHSLSASHMETTHKLTKTQGRALAHQFASLEHNLVPTSDQVLPFAKVVSRVLQAMTV
AHEDVNLHFTKDIDVAGLADGTHLYMALLEETAKLQAQVLMVPSSTPPVQLCLMKGKTNNDONDIRA
REDEWVWYKELCGMPSHQLTLTQQLQVLLDVLYLETSHSDVEGKEPQDKLRLRFGPSRNPFFKYNH
VSESSKYNPHTKIDVAGLADGTHLYMALLEETAKLQAQVLMVPSSTPPVQYSEAVOLRDRDVEVYKTCQELQRLHAEIQDKSR
```

Generating ten models and calculating the DOPE score for each:

>> Summary of successfully produced models:			
Filename	molpdf	DOPE score	GA341 score
THC05.B99990001.pdb	14039.76172	-61091.00781	1.00000
THC05.B99990002.pdb	13689.39844	-61583.10547	1.00000
THC05.B99990003.pdb	14012.65527	-60554.16016	1.00000
THC05.B99990004.pdb	15351.11328	-60020.52344	1.00000
THC05.B99990005.pdb	15207.13086	-59362.60547	1.00000
THC05.B99990006.pdb	13705.48730	-62052.34766	1.00000
THC05.B99990007.pdb	13694.52148	-60961.15234	1.00000
THC05.B99990008.pdb	13732.67090	-60762.07031	1.00000
THC05.B99990009.pdb	15065.17383	-59361.82813	1.00000
THC05.B99990010.pdb	13841.96680	-61761.55859	1.00000
Total CPU time [seconds]		:	812.53

Best model: THC05.B99990006.pdb

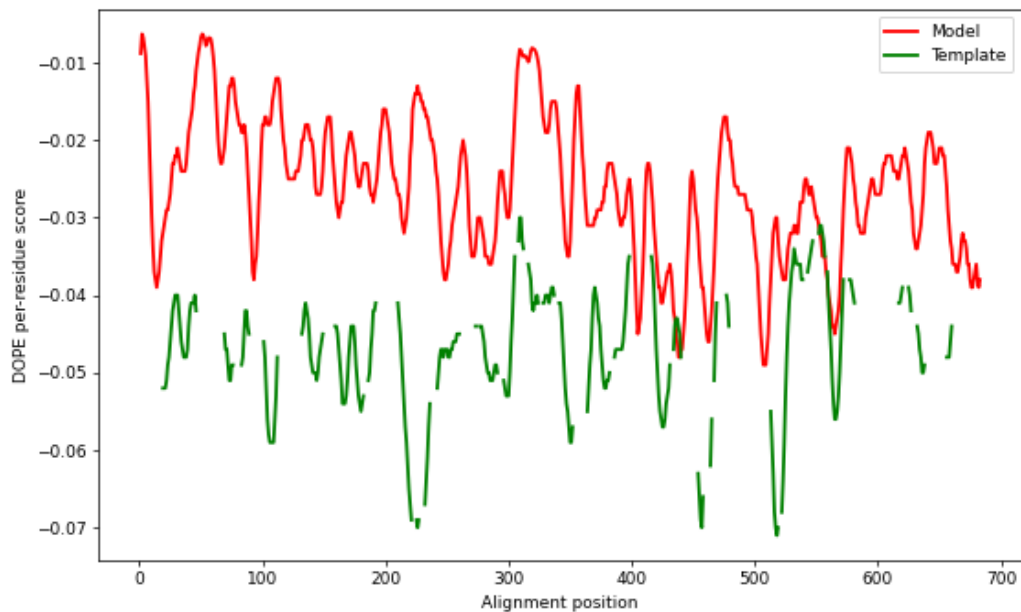
Attach a plot comparing the DOPE score profile of selected templates and the best model (Advanced).



Attach a plot comparing the DOPE score profile of the best model generated from basic and advanced modeling and write about their structural differences.

DOPE score profile of the best model generated from basic modeling:

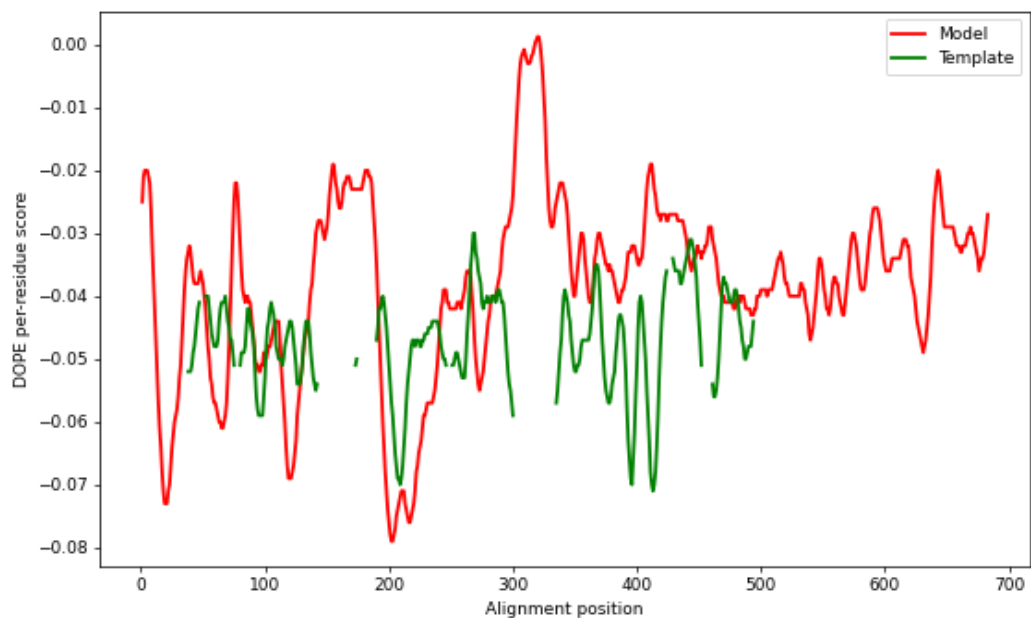
```
>> Summary of successfully produced models:
Filename                               molpdf      DOPE score    GA341 score
-----
THC05.B99990001.pdb                   5361.35303   -42149.51953   0.11013
THC05.B99990002.pdb                   5960.03027   -39069.76953   0.05055
THC05.B99990003.pdb                   5831.55713   -42991.85547   0.08512
THC05.B99990004.pdb                   5287.98389   -43196.62500   0.20044
THC05.B99990005.pdb                   5772.12646   -42612.83984   0.11936
THC05.B99990006.pdb                   5198.93213   -43242.58594   0.16481
THC05.B99990007.pdb                   5998.61865   -42732.57422   0.06710
THC05.B99990008.pdb                   6179.66113   -41441.01563   0.13697
THC05.B99990009.pdb                   5930.36084   -40859.38672   0.06257
THC05.B99990010.pdb                   5441.37500   -43623.36719   0.08697
Total CPU time [seconds]                :      650.66
```



DOPE score profile of the best model generated from advanced modeling:

```
>> Summary of successfully produced models:
Filename                               molpdf      DOPE score    GA341 score
-----
THC05.B99990001.pdb                   14039.76172  -61091.00781  1.00000
THC05.B99990002.pdb                   13689.39844  -61583.10547  1.00000
THC05.B99990003.pdb                   14012.65527  -60554.16016  1.00000
THC05.B99990004.pdb                   15351.11328  -60020.52344  1.00000
THC05.B99990005.pdb                   15207.13086  -59362.60547  1.00000
THC05.B99990006.pdb                   13705.48730  -62052.34766  1.00000
THC05.B99990007.pdb                   13694.52148  -60961.15234  1.00000
THC05.B99990008.pdb                   13732.67090  -60762.07031  1.00000
THC05.B99990009.pdb                   15065.17383  -59361.82813  1.00000
THC05.B99990010.pdb                   13841.96680  -61761.55859  1.00000

Total CPU time [seconds]                :      812.53
```



The model using best modeling is less superimposed than advanced modeling.

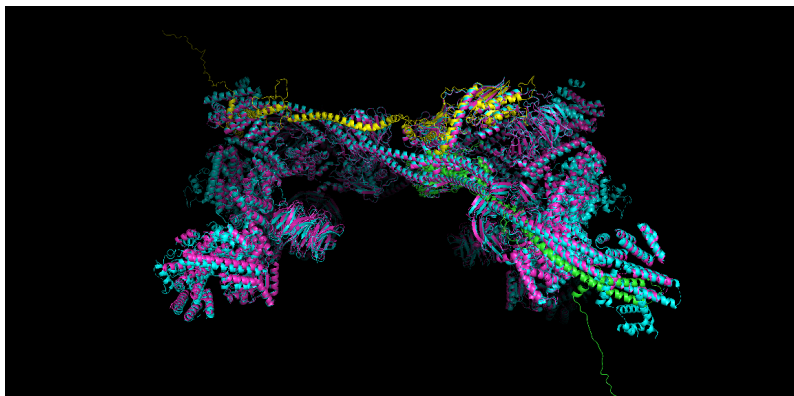
Align all the selected templates with the best model obtained using basic and advanced modeling methods and show their RMSD differences.

The RMSD difference of advanced modeling is greater than that of basic modeling.

```
PyMOL>align 7zn1, 7apk
Match: read scoring matrix.
Match: assigning 11173 x 9967 pairwise scores.
MatchAlign: aligning residues (11173 vs 9967)...
MatchAlign: score 49087.000
ExecutiveAlign: 71921 atoms aligned.
ExecutiveRMS: 1740 atoms rejected during cycle 1 (RMSD=6.32).
ExecutiveRMS: 3646 atoms rejected during cycle 2 (RMSD=2.55).
ExecutiveRMS: 4682 atoms rejected during cycle 3 (RMSD=1.90).
ExecutiveRMS: 4176 atoms rejected during cycle 4 (RMSD=1.55).
ExecutiveRMS: 4822 atoms rejected during cycle 5 (RMSD=1.30).
Executive: RMSD = 1.044 (52855 to 52855 atoms)
```

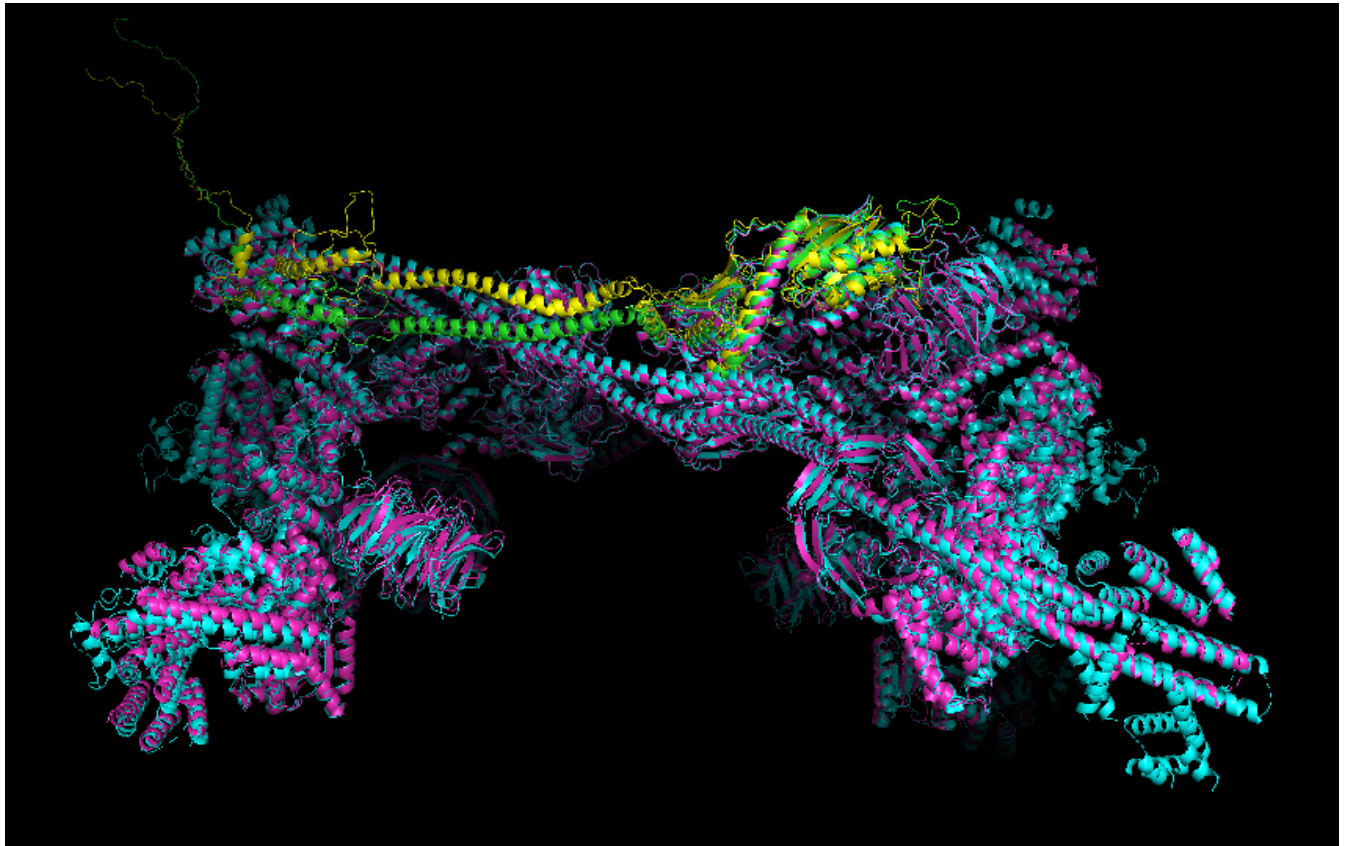
Align all the selected templates (7zn1) with the best model(THC05.B99) obtained using basic and advanced modeling methods

```
PyMOL>align THC05.B99990006, 7zn1
Match: read scoring matrix.
Match: assigning 683 x 11173 pairwise scores.
MatchAlign: aligning residues (683 vs 11173)...
MatchAlign: score 2759.000
ExecutiveAlign: 4145 atoms aligned.
ExecutiveRMS: 216 atoms rejected during cycle 1 (RMSD=19.80).
ExecutiveRMS: 237 atoms rejected during cycle 2 (RMSD=16.05).
ExecutiveRMS: 337 atoms rejected during cycle 3 (RMSD=13.79).
ExecutiveRMS: 249 atoms rejected during cycle 4 (RMSD=9.72).
ExecutiveRMS: 127 atoms rejected during cycle 5 (RMSD=6.87).
Executive: RMSD = 5.645 (2979 to 2979 atoms)
```



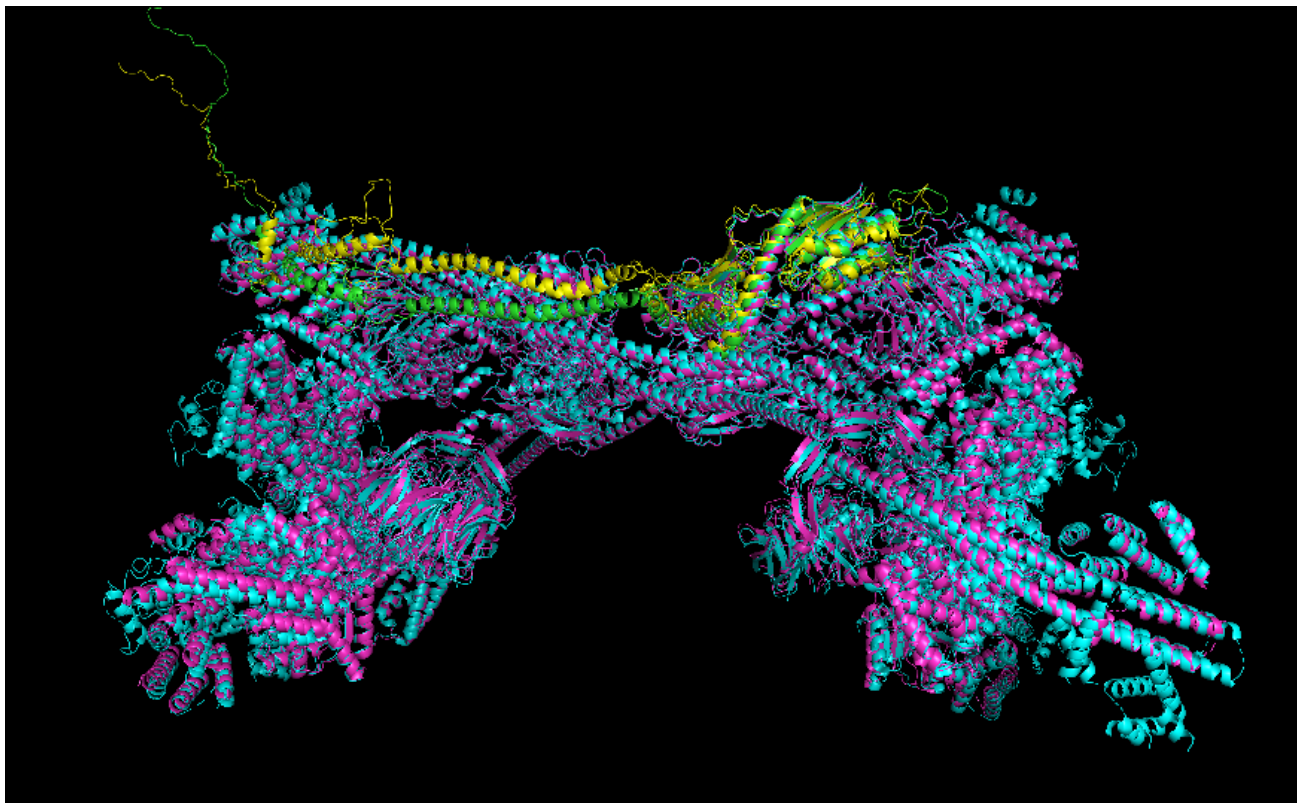
Align all the selected templates (7apk) with the best model(THC05.B99990006) obtained using an advanced modeling method.

```
PyMOL>align THC05.B99990006, 7apk
Match: read scoring matrix.
Match: assigning 683 x 9967 pairwise scores.
MatchAlign: aligning residues (683 vs 9967)...
MatchAlign: score 2752.500
ExecutiveAlign: 4214 atoms aligned.
ExecutiveRMS: 227 atoms rejected during cycle 1 (RMSD=19.26).
ExecutiveRMS: 256 atoms rejected during cycle 2 (RMSD=16.03).
ExecutiveRMS: 311 atoms rejected during cycle 3 (RMSD=13.65).
ExecutiveRMS: 236 atoms rejected during cycle 4 (RMSD=9.69).
ExecutiveRMS: 176 atoms rejected during cycle 5 (RMSD=7.16).
Executive: RMSD = 5.748 (3008 to 3008 atoms)
```



Align all the selected templates(7apk,7znl) with the best model(THC05.B99990010) obtained using basic modeling method

```
PyMOL>align THC05.B99990010, 7apk
Match: read scoring matrix.
Match: assigning 683 x 9967 pairwise scores.
MatchAlign: aligning residues (683 vs 9967)...
MatchAlign: score 2752.500
ExecutiveAlign: 4214 atoms aligned.
ExecutiveRMS: 247 atoms rejected during cycle 1 (RMSD=18.20).
ExecutiveRMS: 301 atoms rejected during cycle 2 (RMSD=14.24).
ExecutiveRMS: 380 atoms rejected during cycle 3 (RMSD=11.41).
ExecutiveRMS: 256 atoms rejected during cycle 4 (RMSD=6.92).
ExecutiveRMS: 152 atoms rejected during cycle 5 (RMSD=3.98).
Executive: RMSD = 2.502 (2878 to 2878 atoms)
```



```
PyMOL>align THC05.B99990010, 7zn1
Match: read scoring matrix.
Match: assigning 683 x 11173 pairwise scores.
MatchAlign: aligning residues (683 vs 11173)...
MatchAlign: score 2759.000
ExecutiveAlign: 4145 atoms aligned.
ExecutiveRMS: 239 atoms rejected during cycle 1 (RMSD=18.94).
ExecutiveRMS: 287 atoms rejected during cycle 2 (RMSD=14.35).
ExecutiveRMS: 381 atoms rejected during cycle 3 (RMSD=11.50).
ExecutiveRMS: 248 atoms rejected during cycle 4 (RMSD=6.96).
ExecutiveRMS: 142 atoms rejected during cycle 5 (RMSD=3.65).
Executive: RMSD = 2.328 (2848 to 2848 atoms)
```

