modeller_automodel

November 8, 2022

[11]:
```
from modeller import *
from modeller.automodel import *
from Bio import SeqIO
from Bio.SeqUtils import seq1
import os

#from modeller import soap_protein_od


pdb_code = '4wkq'
ligand_name_list = ['IRE-A','AAA-A']
output_path = "/Users/rcheeter/Desktop/MODELLER-OUTPUT-GLY"
pdb_file_path = "/users/rcheeter/Desktop/4wkq.pdb"


# FIX GET OF PDB CODE -- MUST EITHER BE IN HEADER LINE OR FIND ANOTHER WAY TO␣
 ↪GET PDB CODE
# CANNOT RELY ON HEADER LINE
# ALSO FIX IN PGN PROGRAM (GET PDB AND GET LIGPLOT ESPECIALLY)

def get_alignment_pir(pdb_code,pdb_file_path,ligand_name_list,output_path):
    print(f'\n[running get_alignment_pir for [{pdb_file_path}]]')

    # import PDB file
    try:
        with open(pdb_file_path,'r') as pdb_file:
            pdb = pdb_file.readlines()
        n = len(pdb)
        if n==0:
            raise Exception(f'  ERROR: error opening PDB file; invalid PDB file␣
 ↪path')
    except:
        raise Exception(f'  ERROR: error opening PDB file: [{pdb_file_path}];␣
 ↪invalid PDB file path')

    if type(output_path)!=str or ' ' in output_path:
```

```
        raise Exception(f'  ERROR: error reading output path: [{output_path}];␣
↪must be a string without spaces')
    try:
        if not os.path.exists(output_path):
            os.makedirs(output_path)
            print(f'  NOTE: creating output path: [{output_path}]')
    except:
        raise Exception(f'  ERROR: error reading output path: [{output_path}]')

    # save original PDB file
    original_pdb_output_path = f'{output_path}/{pdb_code}_original'
    try:
        if not os.path.exists(original_pdb_output_path):
            os.makedirs(original_pdb_output_path)
            print(f'  NOTE: creating original PDB output path:␣
↪[{original_pdb_output_path}]')
    except:
        raise Exception(f'  ERROR: error creating original PDB output path:␣
↪[{original_pdb_output_path}]')

    original_pdb_file_path = f'{original_pdb_output_path}/{pdb_code}.pdb'
    open(original_pdb_file_path,'w').write(''.join(pdb))
    print(f'  NOTE: saving original PDB file to original PDB output path:␣
↪[{original_pdb_file_path}]')

    # modify PDB file
    residue_name_list =␣
↪['ALA','VAL','ILE','LEU','MET','PHE','TYR','TRP','CYS','GLY',
                         ␣
↪'PRO','SER','THR','ASN','GLN','ARG','HIS','LYS','ASP','GLU']
    ligand_name_pdb_set = []

    n = len(pdb)
    i = 0
    while i<n:
        line = pdb[i]
        if line[0:6].upper()=='HETATM':
            ligand_name = f"{line[17:20].replace(' ','').upper()}-{line[21].
↪upper()}"
            ligand_name_pdb_set.append(ligand_name)

        if (line[0:6].upper() in ['ATOM','ANISOU']) and (line[17:20].replace('␣
↪' ,'').upper() not in residue_name_list):
            atom_name = line[12:16].replace(' ','').upper()
            residue_number = int(line[22:26])
            residue_name = line[17:20].replace(' ','').upper()
```

```
            chain_id = line[21].upper()

            # residue_name_1L = seq1(residue_name_3L)
            # removed_residues.
↪append([residue_number,residue_name_3L,residue_name_1L])

            print(f'  WARNING: unrecognized ATOM/ANISOU residue: atom␣
↪[{atom_name}] in residue [{residue_name}-{chain_id} #{residue_number}];␣
↪removing atom')
            pdb.remove(line)
            i -= 1
            n -= 1

        elif (line[0:6].upper()=='HETATM') and ((f"{line[17:20].replace(' ','').
↪upper()}-{line[21].upper()}") not in ligand_name_list):
            atom_name = line[12:16].replace(' ','').upper()
            residue_number = int(line[22:26])
            residue_name = line[17:20].replace(' ','').upper()
            chain_id = line[21].upper()

            # residue_name_1L = seq1(residue_name_3L)
            # removed_residues.
↪append([residue_number,residue_name_3L,residue_name_1L])

            print(f'  WARNING: unrecognized HETATM residue: atom [{atom_name}]␣
↪in residue [{residue_name}-{chain_id} #{residue_number}]; removing atom')
            pdb.remove(line)
            i -= 1
            n -= 1

        i += 1

    # removed_residues = [list(residue) for residue in set(tuple(residue) for␣
↪residue in removed_residues)]
    # print(removed_residues)

    ligand_name_pdb_set = list(set(ligand_name_pdb_set))
    ligand_name_list = list(set(ligand_name_list))
    ligand_name_difference =␣
↪list(set(ligand_name_list)-set(ligand_name_pdb_set))

    if len(ligand_name_difference)!=0:
        ligand_name_difference = ', '.join(ligand_name_difference)
        print(f'  WARNING: ligand name(s) not found in PDB:␣
↪[{ligand_name_difference}]; removing from ligand name list')
```

```
    ligand_name_list = [ligand_name for ligand_name in ligand_name_list if␣
↪ligand_name not in ligand_name_difference]

    # save modified PDB file
    modified_pdb_output_path = f'{output_path}/{pdb_code}_modified'
    try:
        if not os.path.exists(modified_pdb_output_path):
            os.makedirs(modified_pdb_output_path)
            print(f'  NOTE: creating modified PDB output path:␣
↪[{modified_pdb_output_path}]')
    except:
        raise Exception(f'  ERROR: error creating modified PDB output path:␣
↪[{modified_pdb_output_path}]')

    modified_pdb_file_path = f'{modified_pdb_output_path}/{pdb_code}_mod.pdb'
    open(modified_pdb_file_path,'w').write(''.join(pdb))
    print(f'  NOTE: saving modified PDB file to modified PDB output path:␣
↪[{modified_pdb_file_path}]')

    # create template and target sequences from PDB file
    template_sequence = ''
    with open(modified_pdb_file_path,'r') as modified_pdb_file:
        for record in SeqIO.parse(modified_pdb_file,'pdb-atom'):
            template_sequence += ('/' + str(record.seq).replace('X','/').
↪replace(' ',''))
        if pdb_code!=(record.id.split(':')[0].lower()):
            raise Exception('  ERROR: error parsing modified PDB file; PDB code␣
↪does not mach original PDB code')

    template_sequence = list(template_sequence)[1:]
    for i in range(1,len(template_sequence)):
        if template_sequence[i]=='/' and (template_sequence[i-1] in ['/','-']):
            template_sequence[i] = '-'
    template_sequence = ''.join(template_sequence)

    for ligand_name in ligand_name_list:
        template_sequence += '/.'
    template_sequence += '*'

    n = 75
    template_sequence = [template_sequence[i:i+n] for i in␣
↪range(0,len(template_sequence),n)]
    template_name = f'{pdb_code}_template'
    template_info = f'structure:{pdb_code}_mod.pdb   :FIRST:@:  END: :::     : ␣
↪   '
```

```
    template_text = [f'>P1;{template_name}',template_info]
    template_text.extend(template_sequence)
    template_text = '\n'.join(template_text)

    target_sequence = template_sequence
    target_name = f'{pdb_code}_target'
    target_info = f'sequence :{target_name}:     : :     : :::     :     '

    target_text = [f'>P1;{target_name}',target_info]
    target_text.extend(target_sequence)
    target_text = '\n'.join(target_text)

    pir_text = template_text + '\n\n' + target_text + '\n'


    pir_file_path = f'{output_path}/alignment.ali'
    open(pir_file_path,'w').write(pir_text)
    print(f'  NOTE: saving PIR sequence alignment to output path:␣
 ↪[{pir_file_path}]')

    return pir_file_path,modified_pdb_file_path,target_name,template_name


    # mutation = [startind,endind,newseq]
# def make_mutation(pir_file_path,mutation,target_name):

#     return pir_file_path


def get_modeller(pir_file_path,target_name,template_name,output_path):

    os.chdir(output_path)
    env = Environ()
    env.io.atom_files_directory = [f'{output_path}/4wkq_modified']
    env.io.hetatm = True

    a = AutoModel(env, alnfile=pir_file_path,
                  knowns=template_name, sequence=target_name,
                  assess_methods=(assess.DOPE,
                                  #soap_protein_od.Scorer(),
                                  assess.GA341))
    a.starting_model = 1
    a.ending_model = 4
    a.make()
```

```
# pir_file_path,modified_pdb_file_path,target_name,template_name =␣
 ↪get_alignment_pir(pdb_code,pdb_file_path,ligand_name_list,output_path)


# make_mutation

get_modeller(pir_file_path,target_name,template_name,output_path)









# PUT TEAM MEETING SLIDES IN FOLDER ON MICROSOFT TEAMS




#         for i in range(1,len(sequence)):
#                 if template_sequence[i]=='/' and (template_sequence[i-1] in␣
 ↪['/','-']):
#                     template_sequence[i] = '-'
#             template_sequence = ''.join(template_sequence)




            # REMOVE / ON LAST CHAIN
            # FIX OUTPUT OF THIS TO BE PROPERLY FORMATTED

#             ali_text.append
#             ali_text.append(f'>P1;{record.id}\n')
#             ali_text.append(f'
#             ali_text.append(template_sequence + '\n\n')

#     n = 75
#     ali_text = [ali_text[i:i+n] for i in range(0, len(ali_text),n)]
```

```
#     ali_text.insert(0,f'>P1;ABCD')
#     ali_text.insert(1,f'asdfadsfadsfasdfasdfadsf')
#     print(ali_text)
#     print(type(ali_text))
#     ali_text = '\n'.join(ali_text)




# hetatm_name_list = ['IRE-A']

# output_path = "/Users/rcheeter/Desktop/PGN_MODELLER_OUTPUT"
# pdb_file_path = f"{output_path}/4wkq.pdb"
# pir_file_path = f"{output_path}/zsample.ali"

# if not os.path.exists(output_path):
#     os.makedirs(output_path)
#     print(f'  NOTE: creating output path directory: [{output_path}]')

# with open(pdb_file_path,'r') as pdb_file,open(pir_file_path,'w') as pir_file:
#         sequences = SeqIO.parse(pdb_file,'pdb-atom')
#         SeqIO.write(sequences,pir_file,'pir')



# ali_text = []
# with open(pdb_file_path, 'r') as pdb_file:
#     for record in SeqIO.parse(pdb_file, 'pdb-atom'):
#         ali_text.append(f'>P1;{record.id}\n')
#         ali_text.append(str(record.seq))

#         print('>' + record.id)
#         print(record.seq)

# open(pir_file_path,'w').write(''.join(ali_text)) # write PDB to a new file in␣
 ↪output path
```

```
rdpdb___459W> Residue type  IRE not recognized. 'AutoModel' model building
              will treat this residue as a rigid body.
              To use real parameters, add the residue type to ${LIB}/restyp.lib,
              its topology to ${LIB}/top_*.lib, and suitable forcefield
              parameters to ${LIB}/par.lib.
fndatmi_285W> Only      296 residues out of      297 contain atoms of type  CA
              (This is usually caused by non-standard residues, such
              as ligands, or by PDB files with missing atoms.)
```

```
fndatmi_285W> Only      296 residues out of      297 contain atoms of type  CA
              (This is usually caused by non-standard residues, such
              as ligands, or by PDB files with missing atoms.)

check_ali___> Checking the sequence-structure alignment.

Implied intrachain target CA(i)-CA(i+1) distances longer than  8.0 angstroms:

ALN_POS  TMPL  RID1  RID2  NAM1  NAM2     DIST
-----------------------------------------------
END OF TABLE
read_to_681_> topology.submodel read from topology file:        3
fndatmi_285W> Only      296 residues out of      297 contain atoms of type  CA
              (This is usually caused by non-standard residues, such
              as ligands, or by PDB files with missing atoms.)
mdtrsr__446W> A potential that relies on one protein is used, yet you have at
              least one known structure available. MDT, not library, potential
is used.
iup2crm_280W> No topology library in memory or assigning a BLK residue.
              Default CHARMM atom type assigned:  CAO -->  CT2
              This message is written only for the first such atom.
31 atoms in HETATM/BLK residues constrained
to protein atoms within 2.30 angstroms
and protein CA atoms within 10.00 angstroms
31 atoms in residues without defined topology
constrained to be rigid bodies
condens_443_> Restraints marked for deletion were removed.
              Total number of restraints before, now:    32379    30238
iupac_m_397W> Atoms were not swapped because of the uncertainty of how to handle
the H atom.
iupac_m_484W> Dihedral still outside +-90:      -90.7454
iupac_m_397W> Atoms were not swapped because of the uncertainty of how to handle
the H atom.
>> Model assessment by DOPE potential
iatmcls_286W> MODEL atom not classified:  SER:OXT  SER
iatmcls_286W> MODEL atom not classified:  GLU:OXT  GLU
iatmcls_286W> MODEL atom not classified:  GLY:OXT  GLY
iatmcls_286W> MODEL atom not classified:  ASP:OXT  ASP
iatmcls_286W> MODEL atom not classified:  GLN:OXT  GLN


>> ENERGY; Differences between the model's features and restraints:
Number of all residues in MODEL                  :      297
Number of all, selected real atoms               :     2397    2397
Number of all, selected pseudo atoms             :        0       0
Number of all static, selected restraints        :    30238   30238
COVALENT_CYS                                     :        F
NONBONDED_SEL_ATOMS                              :        1
```

```
Number of non-bonded pairs (excluding 1-2,1-3,1-4):   475259
Dynamic pairs routine                               : 1, NATM x NATM double loop
Atomic shift for contacts update (UPDATE_DYNAMIC) :    0.390
LENNARD_JONES_SWITCH                                :    6.500   7.500
COULOMB_JONES_SWITCH                                :    6.500   7.500
RESIDUE_SPAN_RANGE                                  :        1    9999
NLOGN_USE                                           :       15
CONTACT_SHELL                                       :   15.000
DYNAMIC_PAIRS,_SPHERE,_COULOMB,_LENNARD,_MODELLER :        T       F       F
F       T
SPHERE_STDV                                         :    0.050
RADII_FACTOR                                        :    0.820
Current energy                                      :      -34417.2500



<< end of ENERGY.
DOPE score               : -34417.250000
>> Model assessment by GA341 potential

Surface library          : /opt/anaconda3/lib/modeller-10.1/modlib/surf5.de
Pair library             : /opt/anaconda3/lib/modeller-10.1/modlib/pair9.de
Chain identifier         : _
% sequence identity      :    99.663002
Sequence length          :           27
Compactness              :     0.043058
Native energy (pair)     :    11.591245
Native energy (surface)  :     3.848097
Native energy (combined) :     5.974587
Z score (pair)           :     0.206631
Z score (surface)        :     0.356791
Z score (combined)       :     0.383440
GA341 score              :     0.351118



>> ENERGY; Differences between the model's features and restraints:
Number of all residues in MODEL                     :      297
Number of all, selected real atoms                  :     2397    2397
Number of all, selected pseudo atoms                :        0       0
Number of all static, selected restraints           :    30238   30238
COVALENT_CYS                                        :        F
NONBONDED_SEL_ATOMS                                 :        1
Number of non-bonded pairs (excluding 1-2,1-3,1-4):     5093
Dynamic pairs routine                               : 2, NATM x NATM cell sorting
Atomic shift for contacts update (UPDATE_DYNAMIC) :    0.390
LENNARD_JONES_SWITCH                                :    6.500   7.500
```

```
COULOMB_JONES_SWITCH                                  :    6.500   7.500
RESIDUE_SPAN_RANGE                                    :        0   99999
NLOGN_USE                                             :       15
CONTACT_SHELL                                         :    4.000
DYNAMIC_PAIRS,_SPHERE,_COULOMB,_LENNARD,_MODELLER     :        T       T       F
F       F
SPHERE_STDV                                           :    0.050
RADII_FACTOR                                          :    0.820
Current energy                                        :        1998.5022
```

Summary of the restraint violations:

```
   NUM     … number of restraints.
   NUMVI   … number of restraints with RVIOL > VIOL_REPORT_CUT[i].
   RVIOL   … relative difference from the best value.
   NUMVP   … number of restraints with -Ln(pdf) > VIOL_REPORT_CUT2[i].
   RMS_1   … RMS(feature, minimally_violated_basis_restraint, NUMB).
   RMS_2   … RMS(feature, best_value, NUMB).
   MOL.PDF … scaled contribution to -Ln(Molecular pdf).
```

#	RESTRAINT_GROUP	NUM	NUMVI	NUMVP	RMS_1	RMS_2	MOL.PDF	S_i
1	Bond length potential :	2414	0	0	0.006	0.006	21.624	1.000
2	Bond angle potential :	3266	0	10	2.106	2.106	282.40	1.000
3	Stereochemical cosine torsion poten:	1570	0	38	47.005	47.005	531.29	1.000
4	Stereochemical improper torsion pot:	967	0	0	1.403	1.403	40.055	1.000
5	Soft-sphere overlap restraints :	5093	0	0	0.002	0.002	1.6673	1.000
6	Lennard-Jones 6-12 potential :	0	0	0	0.000	0.000	0.0000	1.000
7	Coulomb point-point electrostatic p:	0	0	0	0.000	0.000	0.0000	1.000
8	H-bonding potential :	0	0	0	0.000	0.000	0.0000	1.000
9	Distance restraints 1 (CA-CA) :	5955	0	0	0.102	0.102	75.237	1.000
10	Distance restraints 2 (N-O) :	5904	2	13	0.194	0.194	214.09	1.000

```
11 Mainchain Phi dihedral restraints  :       0       0      0   0.000   0.000
0.0000       1.000
12 Mainchain Psi dihedral restraints  :       0       0      0   0.000   0.000
0.0000       1.000
13 Mainchain Omega dihedral restraints:     291       0      5   4.066   4.066
56.740       1.000
14 Sidechain Chi_1 dihedral restraints:     259       0     11  64.754  64.754
81.326       1.000
15 Sidechain Chi_2 dihedral restraints:     206       0      6  67.985  67.985
91.981       1.000
16 Sidechain Chi_3 dihedral restraints:      95       0      1  77.368  77.368
68.856       1.000
17 Sidechain Chi_4 dihedral restraints:      36       0      0  80.572  80.572
23.357       1.000
18 Disulfide distance restraints      :       0       0      0   0.000   0.000
0.0000       1.000
19 Disulfide angle restraints         :       0       0      0   0.000   0.000
0.0000       1.000
20 Disulfide dihedral angle restraints:       0       0      0   0.000   0.000
0.0000       1.000
21 Lower bound distance restraints    :       0       0      0   0.000   0.000
0.0000       1.000
22 Upper bound distance restraints    :       0       0      0   0.000   0.000
0.0000       1.000
23 Distance restraints 3 (SDCH-MNCH)  :    4676       0      0   0.447   0.447
179.72       1.000
24 Sidechain Chi_5 dihedral restraints:       0       0      0   0.000   0.000
0.0000       1.000
25 Phi/Psi pair of dihedral restraints:     286       9     22  23.933  35.877
24.359       1.000
26 Distance restraints 4 (SDCH-SDCH)  :    2943       0      3   0.843   0.843
300.41       1.000
27 Distance restraints 5 (X-Y)        :    1370       0      0   0.021   0.021
5.3993       1.000
28 NMR distance restraints 6 (X-Y)    :       0       0      0   0.000   0.000
0.0000       1.000
29 NMR distance restraints 7 (X-Y)    :       0       0      0   0.000   0.000
0.0000       1.000
30 Minimal distance restraints        :       0       0      0   0.000   0.000
0.0000       1.000
31 Non-bonded restraints              :       0       0      0   0.000   0.000
0.0000       1.000
32 Atomic accessibility restraints    :       0       0      0   0.000   0.000
0.0000       1.000
33 Atomic density restraints          :       0       0      0   0.000   0.000
0.0000       1.000
34 Absolute position restraints       :       0       0      0   0.000   0.000
0.0000       1.000
```

```
35 Dihedral angle difference restraint:       0       0      0   0.000   0.000
0.0000       1.000
36 GBSA implicit solvent potential    :       0       0      0   0.000   0.000
0.0000       1.000
37 EM density fitting potential       :       0       0      0   0.000   0.000
0.0000       1.000
38 SAXS restraints                    :       0       0      0   0.000   0.000
0.0000       1.000
39 Symmetry restraints                :       0       0      0   0.000   0.000
0.0000       1.000



# Heavy relative violation of each residue is written to: 4wkq_target.V99990001
# The profile is NOT normalized by the number of restraints.
# The profiles are smoothed over a window of residues:    1
# The sum of all numbers in the file:   28750.2832



List of the violated restraints:
   A restraint is violated when the relative difference
   from the best value (RVIOL) is larger than CUTOFF.

   ICSR   … index of a restraint in the current set.
   RESNO  … residue numbers of the first two atoms.
   ATM    … IUPAC atom names of the first two atoms.
   FEAT   … the value of the feature in the model.
   restr  … the mean of the basis restraint with the smallest
              difference from the model (local minimum).
   viol   … difference from the local minimum.
   rviol  … relative difference from the local minimum.
   RESTR  … the best value (global minimum).
   VIOL   … difference from the best value.
   RVIOL  … relative difference from the best value.


-------------------------------------------------------------------------------
-----------------

Feature 10                            : Distance restraints 2 (N-O)
List of the RVIOL violations larger than   :       4.5000

    #   ICSR  RESNO1/2 ATM1/2   INDATM1/2    FEAT   restr    viol   rviol
RESTR    VIOL   RVIOL
    1  17294 107I 280Q N   O     827 2240    9.49    6.93    2.56    4.78
6.93    2.56    4.78
```

```
--------------------------------------------------------------------------------
----------------

Feature 25                            : Phi/Psi pair of dihedral restraints
List of the RVIOL violations larger than   :       6.5000

    #   ICSR  RESNO1/2 ATM1/2   INDATM1/2    FEAT   restr    viol   rviol
RESTR    VIOL   RVIOL
    1   8222   5A   6P C   N      30   32  -69.16  -58.70   13.23    1.06
-64.50  169.66   12.91
    1          6P   6P N   CA     32   33  -22.40  -30.50
147.20
    2   8252  37P  38E C   N     288  290   12.39  -63.60  116.31   13.65
-63.60  116.31   13.65
    2         38E  38E N   CA    290  291 -128.36  -40.30
-40.30
    3   8345 134R 135D C   N    1061 1063 -109.86  -63.30   65.38    7.41
-63.30   65.38    7.41
    3        135D 135D N   CA   1063 1064    5.91  -40.00
-40.00
    4   8422 211K 212P C   N    1678 1680  -48.78  -58.70   44.70    2.97
-64.50  139.60   11.26
    4        212P 212P N   CA   1680 1681  -74.09  -30.50
147.20
    5   8423 212P 213Y C   N    1685 1687 -107.14  -98.40   35.50    3.63
55.90  171.90   16.06
    5        213Y 213Y N   CA   1687 1688   93.99  128.40
39.50
    6   8424 213Y 214D C   N    1697 1699  -57.85  -63.30   16.20    1.75
-70.90  155.00    8.29
    6        214D 214D N   CA   1699 1700  -55.26  -40.00
150.30
    7   8425 214D 215G C   N    1705 1707  -72.75  -62.40   12.80    1.89
82.20  160.59   12.01
    7        215G 215G N   CA   1707 1708  -33.67  -41.20
8.50
    8   8490 279I 280Q C   N    2230 2232  -61.84  -63.80   23.01    3.28
-73.00  156.47   10.46
    8        280Q 280Q N   CA   2232 2233  -63.23  -40.30
140.70
    9   8491 280Q 281G C   N    2239 2241  -61.71  -62.40   17.08    2.58
82.20  147.57   10.90
    9        281G 281G N   CA   2241 2242  -24.13  -41.20
8.50


report______> Distribution of short non-bonded contacts:
```

```
DISTANCE1:  0.00 2.10 2.20 2.30 2.40 2.50 2.60 2.70 2.80 2.90 3.00 3.10 3.20
3.30 3.40
DISTANCE2:  2.10 2.20 2.30 2.40 2.50 2.60 2.70 2.80 2.90 3.00 3.10 3.20 3.30
3.40 3.50
FREQUENCY:     0    0    0    0    0   11   21  101  168  225  216  282  303
385  389


<< end of ENERGY.
>> Model assessment by DOPE potential
iatmcls_286W> MODEL atom not classified:  SER:OXT  SER
iatmcls_286W> MODEL atom not classified:  GLU:OXT  GLU
iatmcls_286W> MODEL atom not classified:  GLY:OXT  GLY
iatmcls_286W> MODEL atom not classified:  ASP:OXT  ASP
iatmcls_286W> MODEL atom not classified:  GLN:OXT  GLN


>> ENERGY; Differences between the model's features and restraints:
Number of all residues in MODEL                   :      297
Number of all, selected real atoms                :     2397    2397
Number of all, selected pseudo atoms              :        0       0
Number of all static, selected restraints         :    30238   30238
COVALENT_CYS                                      :        F
NONBONDED_SEL_ATOMS                               :        1
Number of non-bonded pairs (excluding 1-2,1-3,1-4):   475940
Dynamic pairs routine                             : 1, NATM x NATM double loop
Atomic shift for contacts update (UPDATE_DYNAMIC) :    0.390
LENNARD_JONES_SWITCH                              :    6.500   7.500
COULOMB_JONES_SWITCH                              :    6.500   7.500
RESIDUE_SPAN_RANGE                                :        1    9999
NLOGN_USE                                         :       15
CONTACT_SHELL                                     :   15.000
DYNAMIC_PAIRS,_SPHERE,_COULOMB,_LENNARD,_MODELLER :        T       F       F
F       T
SPHERE_STDV                                       :    0.050
RADII_FACTOR                                      :    0.820
Current energy                                    :      -34384.6250




<< end of ENERGY.
DOPE score               : -34384.625000
>> Model assessment by GA341 potential

Surface library          : /opt/anaconda3/lib/modeller-10.1/modlib/surf5.de
Pair library             : /opt/anaconda3/lib/modeller-10.1/modlib/pair9.de
```

```
Chain identifier         : _
% sequence identity      :    99.663002
Sequence length          :           27
Compactness              :     0.042779
Native energy (pair)     :    10.347894
Native energy (surface)  :     2.188346
Native energy (combined) :     4.201102
Z score (pair)           :    -0.029196
Z score (surface)        :    -0.690697
Z score (combined)       :    -0.488425
GA341 score              :     0.644417
```

```
>> ENERGY; Differences between the model's features and restraints:
Number of all residues in MODEL                    :      297
Number of all, selected real atoms                 :     2397    2397
Number of all, selected pseudo atoms               :        0       0
Number of all static, selected restraints          :    30238   30238
COVALENT_CYS                                       :        F
NONBONDED_SEL_ATOMS                                :        1
Number of non-bonded pairs (excluding 1-2,1-3,1-4):     5109
Dynamic pairs routine                              : 2, NATM x NATM cell sorting
Atomic shift for contacts update (UPDATE_DYNAMIC) :    0.390
LENNARD_JONES_SWITCH                               :    6.500   7.500
COULOMB_JONES_SWITCH                               :    6.500   7.500
RESIDUE_SPAN_RANGE                                 :        0   99999
NLOGN_USE                                          :       15
CONTACT_SHELL                                      :    4.000
DYNAMIC_PAIRS,_SPHERE,_COULOMB,_LENNARD,_MODELLER :        T       T       F
F       F
SPHERE_STDV                                        :    0.050
RADII_FACTOR                                       :    0.820
Current energy                                     :        2077.9355
```

```
Summary of the restraint violations:

   NUM     … number of restraints.
   NUMVI   … number of restraints with RVIOL > VIOL_REPORT_CUT[i].
   RVIOL   … relative difference from the best value.
   NUMVP   … number of restraints with -Ln(pdf) > VIOL_REPORT_CUT2[i].
   RMS_1   … RMS(feature, minimally_violated_basis_restraint, NUMB).
   RMS_2   … RMS(feature, best_value, NUMB).
   MOL.PDF … scaled contribution to -Ln(Molecular pdf).
```

```
 #                      RESTRAINT_GROUP      NUM   NUMVI  NUMVP   RMS_1   RMS_2
MOL.PDF     S_i
-------------------------------------------------------------------------------
----------------------
 1 Bond length potential              :    2414       0      1   0.006   0.006
23.649       1.000
 2 Bond angle potential               :    3266       2      8   2.150   2.150
296.68       1.000
 3 Stereochemical cosine torsion poten:    1570       0     39  47.222  47.222
535.37       1.000
 4 Stereochemical improper torsion pot:     967       1      1   2.451   2.451
105.88       1.000
 5 Soft-sphere overlap restraints     :    5109       0      2   0.005   0.005
13.123       1.000
 6 Lennard-Jones 6-12 potential       :       0       0      0   0.000   0.000
0.0000       1.000
 7 Coulomb point-point electrostatic p:       0       0      0   0.000   0.000
0.0000       1.000
 8 H-bonding potential                :       0       0      0   0.000   0.000
0.0000       1.000
 9 Distance restraints 1 (CA-CA)      :    5955       0      0   0.093   0.093
65.978       1.000
10 Distance restraints 2 (N-O)        :    5904       0     10   0.164   0.164
170.73       1.000
11 Mainchain Phi dihedral restraints  :       0       0      0   0.000   0.000
0.0000       1.000
12 Mainchain Psi dihedral restraints  :       0       0      0   0.000   0.000
0.0000       1.000
13 Mainchain Omega dihedral restraints:     291       1      3   4.267   4.267
62.475       1.000
14 Sidechain Chi_1 dihedral restraints:     259       0     13  58.896  58.896
58.589       1.000
15 Sidechain Chi_2 dihedral restraints:     206       0      6  61.224  61.224
75.400       1.000
16 Sidechain Chi_3 dihedral restraints:      95       0      0  65.848  65.848
65.515       1.000
17 Sidechain Chi_4 dihedral restraints:      36       0      0  97.201  97.201
22.955       1.000
18 Disulfide distance restraints      :       0       0      0   0.000   0.000
0.0000       1.000
19 Disulfide angle restraints         :       0       0      0   0.000   0.000
0.0000       1.000
20 Disulfide dihedral angle restraints:       0       0      0   0.000   0.000
0.0000       1.000
21 Lower bound distance restraints    :       0       0      0   0.000   0.000
0.0000       1.000
22 Upper bound distance restraints    :       0       0      0   0.000   0.000
```

```
0.0000       1.000
23 Distance restraints 3 (SDCH-MNCH)  :    4676       0      1   0.471   0.471
200.71       1.000
24 Sidechain Chi_5 dihedral restraints:       0       0      0   0.000   0.000
0.0000       1.000
25 Phi/Psi pair of dihedral restraints:     286       6     22  23.705  26.527
75.218       1.000
26 Distance restraints 4 (SDCH-SDCH)  :    2943       0      9   0.838   0.838
300.55       1.000
27 Distance restraints 5 (X-Y)        :    1370       0      0   0.022   0.022
5.1001       1.000
28 NMR distance restraints 6 (X-Y)    :       0       0      0   0.000   0.000
0.0000       1.000
29 NMR distance restraints 7 (X-Y)    :       0       0      0   0.000   0.000
0.0000       1.000
30 Minimal distance restraints        :       0       0      0   0.000   0.000
0.0000       1.000
31 Non-bonded restraints              :       0       0      0   0.000   0.000
0.0000       1.000
32 Atomic accessibility restraints    :       0       0      0   0.000   0.000
0.0000       1.000
33 Atomic density restraints          :       0       0      0   0.000   0.000
0.0000       1.000
34 Absolute position restraints       :       0       0      0   0.000   0.000
0.0000       1.000
35 Dihedral angle difference restraint:       0       0      0   0.000   0.000
0.0000       1.000
36 GBSA implicit solvent potential    :       0       0      0   0.000   0.000
0.0000       1.000
37 EM density fitting potential       :       0       0      0   0.000   0.000
0.0000       1.000
38 SAXS restraints                    :       0       0      0   0.000   0.000
0.0000       1.000
39 Symmetry restraints                :       0       0      0   0.000   0.000
0.0000       1.000


# Heavy relative violation of each residue is written to: 4wkq_target.V99990002
# The profile is NOT normalized by the number of restraints.
# The profiles are smoothed over a window of residues:    1
# The sum of all numbers in the file:   27202.0742


List of the violated restraints:
   A restraint is violated when the relative difference
   from the best value (RVIOL) is larger than CUTOFF.
```

```
   ICSR   … index of a restraint in the current set.
   RESNO  … residue numbers of the first two atoms.
   ATM    … IUPAC atom names of the first two atoms.
   FEAT   … the value of the feature in the model.
   restr  … the mean of the basis restraint with the smallest
              difference from the model (local minimum).
   viol   … difference from the local minimum.
   rviol  … relative difference from the local minimum.
   RESTR  … the best value (global minimum).
   VIOL   … difference from the best value.
   RVIOL  … relative difference from the best value.


-------------------------------------------------------------------------------
-----------------

Feature 13                          : Mainchain Omega dihedral restraints
List of the RVIOL violations larger than   :       4.5000

    #   ICSR  RESNO1/2 ATM1/2   INDATM1/2    FEAT   restr    viol   rviol
RESTR    VIOL   RVIOL
    1   8547  45P  45P CA  C     346  350 -156.82 -180.00   23.18    4.63
-180.00   23.18    4.63

-------------------------------------------------------------------------------
-----------------

Feature 25                          : Phi/Psi pair of dihedral restraints
List of the RVIOL violations larger than   :       6.5000

    #   ICSR  RESNO1/2 ATM1/2   INDATM1/2    FEAT   restr    viol   rviol
RESTR    VIOL   RVIOL
    1   8252  37P  38E C   N     288  290    7.77  -63.60  111.70   13.06
-63.60  111.70   13.06
    1         38E  38E N   CA    290  291 -126.23  -40.30
-40.30
    2   8259  44I  45P C   N     343  345 -152.39  -64.50   88.58    9.42
-64.50   88.58    9.42
    2         45P  45P N   CA    345  346  136.21  147.20
147.20
    3   8316 105D 106N C   N     817  819 -114.89  -63.20   87.56    9.46
-63.20   87.56    9.46
    3        106N 106N N   CA    819  820   29.57  -41.10
-41.10
    4   8345 134R 135D C   N    1061 1063 -110.28  -63.30   65.49    7.44
-63.30   65.49    7.44
    4        135D 135D N   CA   1063 1064    5.62  -40.00
```

```
-40.00
    5   8437 226E 227K C   N    1789 1791  -68.06  -62.90   61.69    8.30
-62.90   61.69    8.30
    5        227K 227K N   CA   1791 1792 -102.27  -40.80
-40.80
    6   8438 227K 228G C   N    1798 1800 -115.97  -62.40   54.59    9.27
82.20  166.51    9.51
    6        228G 228G N   CA   1800 1801  -30.70  -41.20
8.50


report______> Distribution of short non-bonded contacts:


DISTANCE1:  0.00 2.10 2.20 2.30 2.40 2.50 2.60 2.70 2.80 2.90 3.00 3.10 3.20
3.30 3.40
DISTANCE2:  2.10 2.20 2.30 2.40 2.50 2.60 2.70 2.80 2.90 3.00 3.10 3.20 3.30
3.40 3.50
FREQUENCY:     0    0    0    0    1   12   27  106  162  251  221  256  331
387  390


<< end of ENERGY.
iupac_m_397W> Atoms were not swapped because of the uncertainty of how to handle
the H atom.
>> Model assessment by DOPE potential
iatmcls_286W> MODEL atom not classified:  SER:OXT  SER
iatmcls_286W> MODEL atom not classified:  GLU:OXT  GLU
iatmcls_286W> MODEL atom not classified:  GLY:OXT  GLY
iatmcls_286W> MODEL atom not classified:  ASP:OXT  ASP
iatmcls_286W> MODEL atom not classified:  GLN:OXT  GLN


>> ENERGY; Differences between the model's features and restraints:
Number of all residues in MODEL                    :      297
Number of all, selected real atoms                 :     2397    2397
Number of all, selected pseudo atoms               :        0       0
Number of all static, selected restraints          :    30238   30238
COVALENT_CYS                                       :        F
NONBONDED_SEL_ATOMS                                :        1
Number of non-bonded pairs (excluding 1-2,1-3,1-4):   474622
Dynamic pairs routine                              : 1, NATM x NATM double loop
Atomic shift for contacts update (UPDATE_DYNAMIC) :    0.390
LENNARD_JONES_SWITCH                               :    6.500   7.500
COULOMB_JONES_SWITCH                               :    6.500   7.500
RESIDUE_SPAN_RANGE                                 :        1    9999
NLOGN_USE                                          :       15
CONTACT_SHELL                                      :   15.000
```

```
DYNAMIC_PAIRS,_SPHERE,_COULOMB,_LENNARD,_MODELLER :        T       F       F
F       T
SPHERE_STDV                                       :    0.050
RADII_FACTOR                                      :    0.820
Current energy                                    :      -34578.4492



<< end of ENERGY.
DOPE score               : -34578.449219
>> Model assessment by GA341 potential

Surface library          : /opt/anaconda3/lib/modeller-10.1/modlib/surf5.de
Pair library             : /opt/anaconda3/lib/modeller-10.1/modlib/pair9.de
Chain identifier         : _
% sequence identity      :    99.663002
Sequence length          :           27
Compactness              :     0.042669
Native energy (pair)     :    12.094317
Native energy (surface)  :     2.556446
Native energy (combined) :     4.947094
Z score (pair)           :     0.153613
Z score (surface)        :    -0.500128
Z score (combined)       :    -0.231451
GA341 score              :     0.550490



>> ENERGY; Differences between the model's features and restraints:
Number of all residues in MODEL                   :      297
Number of all, selected real atoms                :     2397    2397
Number of all, selected pseudo atoms              :        0       0
Number of all static, selected restraints         :    30238   30238
COVALENT_CYS                                      :        F
NONBONDED_SEL_ATOMS                               :        1
Number of non-bonded pairs (excluding 1-2,1-3,1-4):     5093
Dynamic pairs routine                             : 2, NATM x NATM cell sorting
Atomic shift for contacts update (UPDATE_DYNAMIC) :    0.390
LENNARD_JONES_SWITCH                              :    6.500   7.500
COULOMB_JONES_SWITCH                              :    6.500   7.500
RESIDUE_SPAN_RANGE                                :        0   99999
NLOGN_USE                                         :       15
CONTACT_SHELL                                     :    4.000
DYNAMIC_PAIRS,_SPHERE,_COULOMB,_LENNARD,_MODELLER :        T       T       F
F       F
SPHERE_STDV                                       :    0.050
RADII_FACTOR                                      :    0.820
```

Current energy : 1821.2335

Summary of the restraint violations:

NUM … number of restraints.
NUMVI … number of restraints with RVIOL > VIOL_REPORT_CUT[i].
RVIOL … relative difference from the best value.
NUMVP … number of restraints with -Ln(pdf) > VIOL_REPORT_CUT2[i].
RMS_1 … RMS(feature, minimally_violated_basis_restraint, NUMB).
RMS_2 … RMS(feature, best_value, NUMB).
MOL.PDF … scaled contribution to -Ln(Molecular pdf).

#	RESTRAINT_GROUP	NUM	NUMVI	NUMVP	RMS_1	RMS_2	MOL.PDF	S_i
1	Bond length potential :	2414	0	0	0.005	0.005	19.511	1.000
2	Bond angle potential :	3266	0	9	2.032	2.032	265.06	1.000
3	Stereochemical cosine torsion poten:	1570	0	39	47.486	47.486	535.45	1.000
4	Stereochemical improper torsion pot:	967	0	0	1.379	1.379	37.909	1.000
5	Soft-sphere overlap restraints :	5093	0	0	0.001	0.001	1.2440	1.000
6	Lennard-Jones 6-12 potential :	0	0	0	0.000	0.000	0.0000	1.000
7	Coulomb point-point electrostatic p:	0	0	0	0.000	0.000	0.0000	1.000
8	H-bonding potential :	0	0	0	0.000	0.000	0.0000	1.000
9	Distance restraints 1 (CA-CA) :	5955	0	0	0.090	0.090	61.019	1.000
10	Distance restraints 2 (N-O) :	5904	0	3	0.140	0.140	119.85	1.000
11	Mainchain Phi dihedral restraints :	0	0	0	0.000	0.000	0.0000	1.000
12	Mainchain Psi dihedral restraints :	0	0	0	0.000	0.000	0.0000	1.000
13	Mainchain Omega dihedral restraints:	291	0	2	3.919	3.919	52.719	1.000
14	Sidechain Chi_1 dihedral restraints:	259	0	8	64.641	64.641	73.416	1.000

```
15 Sidechain Chi_2 dihedral restraints:     206       0      2  60.379  60.379
76.586       1.000
16 Sidechain Chi_3 dihedral restraints:      95       0      1  80.638  80.638
75.979       1.000
17 Sidechain Chi_4 dihedral restraints:      36       0      0 100.316 100.316
24.344       1.000
18 Disulfide distance restraints      :       0       0      0   0.000   0.000
0.0000       1.000
19 Disulfide angle restraints         :       0       0      0   0.000   0.000
0.0000       1.000
20 Disulfide dihedral angle restraints:       0       0      0   0.000   0.000
0.0000       1.000
21 Lower bound distance restraints    :       0       0      0   0.000   0.000
0.0000       1.000
22 Upper bound distance restraints    :       0       0      0   0.000   0.000
0.0000       1.000
23 Distance restraints 3 (SDCH-MNCH)  :    4676       0      0   0.456   0.456
186.75       1.000
24 Sidechain Chi_5 dihedral restraints:       0       0      0   0.000   0.000
0.0000       1.000
25 Phi/Psi pair of dihedral restraints:     286       4     18  20.183  26.888
-20.780       1.000
26 Distance restraints 4 (SDCH-SDCH)  :    2943       0      5   0.850   0.850
307.84       1.000
27 Distance restraints 5 (X-Y)        :    1370       0      0   0.020   0.020
4.3427       1.000
28 NMR distance restraints 6 (X-Y)    :       0       0      0   0.000   0.000
0.0000       1.000
29 NMR distance restraints 7 (X-Y)    :       0       0      0   0.000   0.000
0.0000       1.000
30 Minimal distance restraints        :       0       0      0   0.000   0.000
0.0000       1.000
31 Non-bonded restraints              :       0       0      0   0.000   0.000
0.0000       1.000
32 Atomic accessibility restraints    :       0       0      0   0.000   0.000
0.0000       1.000
33 Atomic density restraints          :       0       0      0   0.000   0.000
0.0000       1.000
34 Absolute position restraints       :       0       0      0   0.000   0.000
0.0000       1.000
35 Dihedral angle difference restraint:       0       0      0   0.000   0.000
0.0000       1.000
36 GBSA implicit solvent potential    :       0       0      0   0.000   0.000
0.0000       1.000
37 EM density fitting potential       :       0       0      0   0.000   0.000
0.0000       1.000
38 SAXS restraints                    :       0       0      0   0.000   0.000
0.0000       1.000
```

```
39 Symmetry restraints                 :       0       0      0   0.000   0.000
0.0000       1.000



# Heavy relative violation of each residue is written to: 4wkq_target.V99990003
# The profile is NOT normalized by the number of restraints.
# The profiles are smoothed over a window of residues:    1
# The sum of all numbers in the file:   27548.9922



List of the violated restraints:
   A restraint is violated when the relative difference
   from the best value (RVIOL) is larger than CUTOFF.

   ICSR   … index of a restraint in the current set.
   RESNO  … residue numbers of the first two atoms.
   ATM    … IUPAC atom names of the first two atoms.
   FEAT   … the value of the feature in the model.
   restr  … the mean of the basis restraint with the smallest
              difference from the model (local minimum).
   viol   … difference from the local minimum.
   rviol  … relative difference from the local minimum.
   RESTR  … the best value (global minimum).
   VIOL   … difference from the best value.
   RVIOL  … relative difference from the best value.


--------------------------------------------------------------------------------
-----------------

Feature 25                             : Phi/Psi pair of dihedral restraints
List of the RVIOL violations larger than   :       6.5000

    #   ICSR  RESNO1/2 ATM1/2   INDATM1/2    FEAT   restr    viol   rviol
RESTR    VIOL   RVIOL
    1   8252  37P  38E C   N     288  290  -63.64  -69.30   11.69    0.70
-63.60  172.57   23.01
    1         38E  38E N   CA    290  291  132.27  142.50
-40.30
    2   8253  38E  39G C   N     297  299   64.11   82.20   18.31    1.32
-167.20 -147.64    8.34
    2         39G  39G N   CA    299  300    5.68    8.50
174.60
    3   8345 134R 135D C   N    1061 1063 -110.76  -63.30   66.92    7.57
-63.30   66.92    7.57
    3        135D 135D N   CA   1063 1064    7.18  -40.00
```

```
-40.00
    4   8350 139R 140N C   N    1098 1100  -98.52  -63.20   75.47    8.13
-63.20   75.47    8.13
    4        140N 140N N   CA   1100 1101   25.60  -41.10
-41.10


report______> Distribution of short non-bonded contacts:


DISTANCE1:  0.00 2.10 2.20 2.30 2.40 2.50 2.60 2.70 2.80 2.90 3.00 3.10 3.20
3.30 3.40
DISTANCE2:  2.10 2.20 2.30 2.40 2.50 2.60 2.70 2.80 2.90 3.00 3.10 3.20 3.30
3.40 3.50
FREQUENCY:     0    0    0    0    0   14   30  103  154  231  220  295  323
354  403


<< end of ENERGY.
>> Model assessment by DOPE potential
iatmcls_286W> MODEL atom not classified:  SER:OXT  SER
iatmcls_286W> MODEL atom not classified:  GLU:OXT  GLU
iatmcls_286W> MODEL atom not classified:  GLY:OXT  GLY
iatmcls_286W> MODEL atom not classified:  ASP:OXT  ASP
iatmcls_286W> MODEL atom not classified:  GLN:OXT  GLN


>> ENERGY; Differences between the model's features and restraints:
Number of all residues in MODEL                    :      297
Number of all, selected real atoms                 :     2397    2397
Number of all, selected pseudo atoms               :        0       0
Number of all static, selected restraints          :    30238   30238
COVALENT_CYS                                       :        F
NONBONDED_SEL_ATOMS                                :        1
Number of non-bonded pairs (excluding 1-2,1-3,1-4):   475455
Dynamic pairs routine                              : 1, NATM x NATM double loop
Atomic shift for contacts update (UPDATE_DYNAMIC) :    0.390
LENNARD_JONES_SWITCH                               :    6.500   7.500
COULOMB_JONES_SWITCH                               :    6.500   7.500
RESIDUE_SPAN_RANGE                                 :        1    9999
NLOGN_USE                                          :       15
CONTACT_SHELL                                      :   15.000
DYNAMIC_PAIRS,_SPHERE,_COULOMB,_LENNARD,_MODELLER :        T       F       F
F       T
SPHERE_STDV                                        :    0.050
RADII_FACTOR                                       :    0.820
Current energy                                     :      -34228.3789
```

```
<< end of ENERGY.
DOPE score               : -34228.378906
>> Model assessment by GA341 potential

Surface library          : /opt/anaconda3/lib/modeller-10.1/modlib/surf5.de
Pair library             : /opt/anaconda3/lib/modeller-10.1/modlib/pair9.de
Chain identifier         : _
% sequence identity      :    99.663002
Sequence length          :           27
Compactness              :     0.043085
Native energy (pair)     :    12.273322
Native energy (surface)  :     2.904532
Native energy (combined) :     5.443948
Z score (pair)           :     0.172798
Z score (surface)        :    -0.337594
Z score (combined)       :    -0.111155
GA341 score              :     0.507988



>> ENERGY; Differences between the model's features and restraints:
Number of all residues in MODEL                   :      297
Number of all, selected real atoms                :     2397    2397
Number of all, selected pseudo atoms              :        0       0
Number of all static, selected restraints         :    30238   30238
COVALENT_CYS                                      :        F
NONBONDED_SEL_ATOMS                               :        1
Number of non-bonded pairs (excluding 1-2,1-3,1-4):     5227
Dynamic pairs routine                             : 2, NATM x NATM cell sorting
Atomic shift for contacts update (UPDATE_DYNAMIC) :    0.390
LENNARD_JONES_SWITCH                              :    6.500   7.500
COULOMB_JONES_SWITCH                              :    6.500   7.500
RESIDUE_SPAN_RANGE                                :        0   99999
NLOGN_USE                                         :       15
CONTACT_SHELL                                     :    4.000
DYNAMIC_PAIRS,_SPHERE,_COULOMB,_LENNARD,_MODELLER :        T       T       F
F       F
SPHERE_STDV                                       :    0.050
RADII_FACTOR                                      :    0.820
Current energy                                    :        1918.2188
```

Summary of the restraint violations:

```
   NUM     … number of restraints.
   NUMVI   … number of restraints with RVIOL > VIOL_REPORT_CUT[i].
   RVIOL   … relative difference from the best value.
   NUMVP   … number of restraints with -Ln(pdf) > VIOL_REPORT_CUT2[i].
   RMS_1   … RMS(feature, minimally_violated_basis_restraint, NUMB).
   RMS_2   … RMS(feature, best_value, NUMB).
   MOL.PDF … scaled contribution to -Ln(Molecular pdf).
```

#	RESTRAINT_GROUP	NUM	NUMVI	NUMVP	RMS_1	RMS_2	MOL.PDF	S_i
1	Bond length potential :	2414	0	0	0.006	0.006	20.569	1.000
2	Bond angle potential :	3266	0	5	2.074	2.074	274.21	1.000
3	Stereochemical cosine torsion poten:	1570	0	41	47.809	47.809	541.92	1.000
4	Stereochemical improper torsion pot:	967	0	0	1.510	1.510	44.840	1.000
5	Soft-sphere overlap restraints :	5227	0	0	0.002	0.002	1.9608	1.000
6	Lennard-Jones 6-12 potential :	0	0	0	0.000	0.000	0.0000	1.000
7	Coulomb point-point electrostatic p:	0	0	0	0.000	0.000	0.0000	1.000
8	H-bonding potential :	0	0	0	0.000	0.000	0.0000	1.000
9	Distance restraints 1 (CA-CA) :	5955	0	0	0.092	0.092	62.114	1.000
10	Distance restraints 2 (N-O) :	5904	0	0	0.138	0.138	118.37	1.000
11	Mainchain Phi dihedral restraints :	0	0	0	0.000	0.000	0.0000	1.000
12	Mainchain Psi dihedral restraints :	0	0	0	0.000	0.000	0.0000	1.000
13	Mainchain Omega dihedral restraints:	291	0	1	4.022	4.022	55.502	1.000
14	Sidechain Chi_1 dihedral restraints:	259	0	13	63.156	63.156	84.821	1.000
15	Sidechain Chi_2 dihedral restraints:	206	0	5	62.658	62.658	94.798	1.000
16	Sidechain Chi_3 dihedral restraints:	95	0	0	69.214	69.214	65.749	1.000
17	Sidechain Chi_4 dihedral restraints:	36	0	0	71.474	71.474	21.631	1.000

```
18 Disulfide distance restraints      :       0       0      0   0.000   0.000
0.0000       1.000
19 Disulfide angle restraints         :       0       0      0   0.000   0.000
0.0000       1.000
20 Disulfide dihedral angle restraints:       0       0      0   0.000   0.000
0.0000       1.000
21 Lower bound distance restraints    :       0       0      0   0.000   0.000
0.0000       1.000
22 Upper bound distance restraints    :       0       0      0   0.000   0.000
0.0000       1.000
23 Distance restraints 3 (SDCH-MNCH)  :    4676       0      0   0.462   0.462
190.92       1.000
24 Sidechain Chi_5 dihedral restraints:       0       0      0   0.000   0.000
0.0000       1.000
25 Phi/Psi pair of dihedral restraints:     286       5     21  24.745  28.634
16.507       1.000
26 Distance restraints 4 (SDCH-SDCH)  :    2943       0      3   0.867   0.867
318.30       1.000
27 Distance restraints 5 (X-Y)        :    1370       0      0   0.023   0.023
6.0131       1.000
28 NMR distance restraints 6 (X-Y)    :       0       0      0   0.000   0.000
0.0000       1.000
29 NMR distance restraints 7 (X-Y)    :       0       0      0   0.000   0.000
0.0000       1.000
30 Minimal distance restraints        :       0       0      0   0.000   0.000
0.0000       1.000
31 Non-bonded restraints              :       0       0      0   0.000   0.000
0.0000       1.000
32 Atomic accessibility restraints    :       0       0      0   0.000   0.000
0.0000       1.000
33 Atomic density restraints          :       0       0      0   0.000   0.000
0.0000       1.000
34 Absolute position restraints       :       0       0      0   0.000   0.000
0.0000       1.000
35 Dihedral angle difference restraint:       0       0      0   0.000   0.000
0.0000       1.000
36 GBSA implicit solvent potential    :       0       0      0   0.000   0.000
0.0000       1.000
37 EM density fitting potential       :       0       0      0   0.000   0.000
0.0000       1.000
38 SAXS restraints                    :       0       0      0   0.000   0.000
0.0000       1.000
39 Symmetry restraints                :       0       0      0   0.000   0.000
0.0000       1.000


# Heavy relative violation of each residue is written to: 4wkq_target.V99990004
```

```
# The profile is NOT normalized by the number of restraints.
# The profiles are smoothed over a window of residues:    1
# The sum of all numbers in the file:   27290.0859



List of the violated restraints:
   A restraint is violated when the relative difference
   from the best value (RVIOL) is larger than CUTOFF.

   ICSR   … index of a restraint in the current set.
   RESNO  … residue numbers of the first two atoms.
   ATM    … IUPAC atom names of the first two atoms.
   FEAT   … the value of the feature in the model.
   restr  … the mean of the basis restraint with the smallest
               difference from the model (local minimum).
   viol   … difference from the local minimum.
   rviol  … relative difference from the local minimum.
   RESTR  … the best value (global minimum).
   VIOL   … difference from the best value.
   RVIOL  … relative difference from the best value.


--------------------------------------------------------------------------------
-----------------

Feature 25                            : Phi/Psi pair of dihedral restraints
List of the RVIOL violations larger than   :       6.5000

    #   ICSR  RESNO1/2 ATM1/2   INDATM1/2    FEAT   restr    viol   rviol
RESTR    VIOL   RVIOL
    1   8222   5A   6P C   N      30   32  -67.85  -58.70   10.98    0.94
-64.50  171.65   12.99
    1          6P   6P N   CA     32   33  -24.42  -30.50
147.20
    2   8223   6P   7N C   N      37   39   64.56  -71.20  144.36    9.10
-119.90 -179.20    8.53
    2          7N   7N N   CA     39   40   93.72  142.80
137.00
    3   8252  37P  38E C   N     288  290   23.41  -69.30  125.08   11.49
-63.60  127.53   15.08
    3         38E  38E N   CA    290  291 -133.53  142.50
-40.30
    4   8316 105D 106N C   N     817  819 -118.07  -63.20   94.11   10.16
-63.20   94.11   10.16
    4        106N 106N N   CA    819  820   35.35  -41.10
-41.10
    5   8345 134R 135D C   N    1061 1063 -110.89  -63.30   66.91    7.58
```

```
-63.30   66.91    7.58
    5        135D 135D N   CA   1063 1064    7.03  -40.00
-40.00


report______> Distribution of short non-bonded contacts:


DISTANCE1:  0.00 2.10 2.20 2.30 2.40 2.50 2.60 2.70 2.80 2.90 3.00 3.10 3.20
3.30 3.40
DISTANCE2:  2.10 2.20 2.30 2.40 2.50 2.60 2.70 2.80 2.90 3.00 3.10 3.20 3.30
3.40 3.50
FREQUENCY:     0    0    0    0    0    9   28  116  163  278  209  290  324
349  418


<< end of ENERGY.

>> Summary of successfully produced models:
Filename                          molpdf     DOPE score    GA341 score
------------------------------------------------------------------------
4wkq_target.B99990001.pdb     1998.50220   -34417.25000        0.35112
4wkq_target.B99990002.pdb     2077.93555   -34384.62500        0.64442
4wkq_target.B99990003.pdb     1821.23352   -34578.44922        0.55049
4wkq_target.B99990004.pdb     1918.21875   -34228.37891        0.50799
```

[ ]:

[ ]:

[ ]:
```python
import re
import numpy as np
string = 'AMGEAPNQALLRILKETEFKKIKVLGSXXXGTVYKGLWIPEGEKVKIPVAIKEXXXXXSPK'
print(str.find('X'))
print(str[27])

# missing_residue_indices = [i.start() for i in re.finditer('X', string)]
# string = list(string)
# string[missing_residue_indices] = '/'

# string = ''.join(string)

sequence = list(sequence.replace('X','/'))
for i in range(1,len(string)):
    if sequence[i]=='/' and (sequence[i-1] in ['/','-']):
        sequence[i] = '-'
```

```
sequence = ''.join(sequence)

# y = re.finditer('X',string)
print(string)
```

[ ]:

[ ]:

[ ]: