Inputs:

pdb: 5HB7, known as Crystal structure of Chaetomium thermophilum Nup53 RRM

seq: fasta sequence of 5HB7:

>5HB7:A|PDBID|CHAIN|SEQUENCE

GPHMPTEVILRGYRNAQHQYAAINHYEQIAGRICEDYPREPPVESRRYKSELRDPAFTHRRALTPEERAKVNRAMSGEHW

VKVTFESAEAADKAVYSSPQLIQGHLVYAEYYKGVPPAQDEAIPD

It has 125 residues.

I run it in the directory /mnt/home/jinchi/bbcontacts-master/5HB7\_test

run

g++ -std=c++0x get\_sorted\_pdb.cpp -o get\_sorted\_pdb.out

./get\_sorted\_pdb.out

to get the distance matrix of the pdb file: pdb\_dis.txt

----------

module load hhblits

hhblits -i 5hb7.fasta -oa3m 5hb7.a3m -all -maxfilt 100000 -realign\_max 100000 -B 100000 -Z 100000 -d /mnt/home/jinchi/hh-downloads/uniprot20/uniprot20\_2013\_03/uniprot20\_2013\_03

The -all option turns off the filtering of the result MSA. Use this option if you want to get all sequences in the significantly similar uniprot20 clusters:

-maxfilt : max number of hits allowed to pass 2nd prefilter (default=20000)

-realign\_max : realign max. hits (default=1000)

“After the search, the most significant

HMMs are realigned using the more accurate Maximum Accuracy (MAC) algorithm”

-B : maximum number of alignments in alignment list (default=500)

-Z maximum number of lines in summary hit list (default=500)

-d <name> database name (e.g. uniprot20\_29Feb2012) (default=)

cp ../Align.pm ../ccmpred ../reformat.pl ../convert\_alignment.py ../HHPaths.pm .

hhfilter -i 5hb7.a3m -o 5hb7.filt.a3m -id 90 -neff 15 -qsc -30

-id [0,100] maximum pairwise sequence identity (%) (def=90)

-neff [1,inf] target diversity of alignment (default=off)

-qsc [0,100] minimum score per column with query (def=-20.0)

../reformat.pl 5hb7.filt.a3m 5hb7.filt.fas -r

~~ssh dev-intel16-k80~~

~~source /mnt/home/jinchi/CCMpred/helper\_modules/bin/activate~~

~~cd /mnt/home/jinchi/bbcontacts-master/5HB7\_test~~

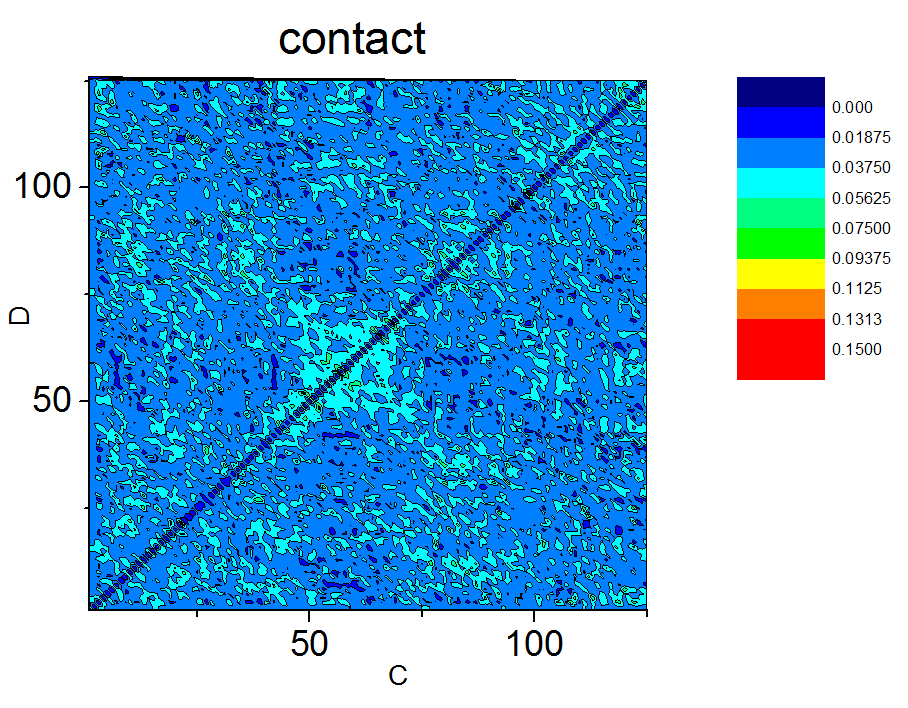
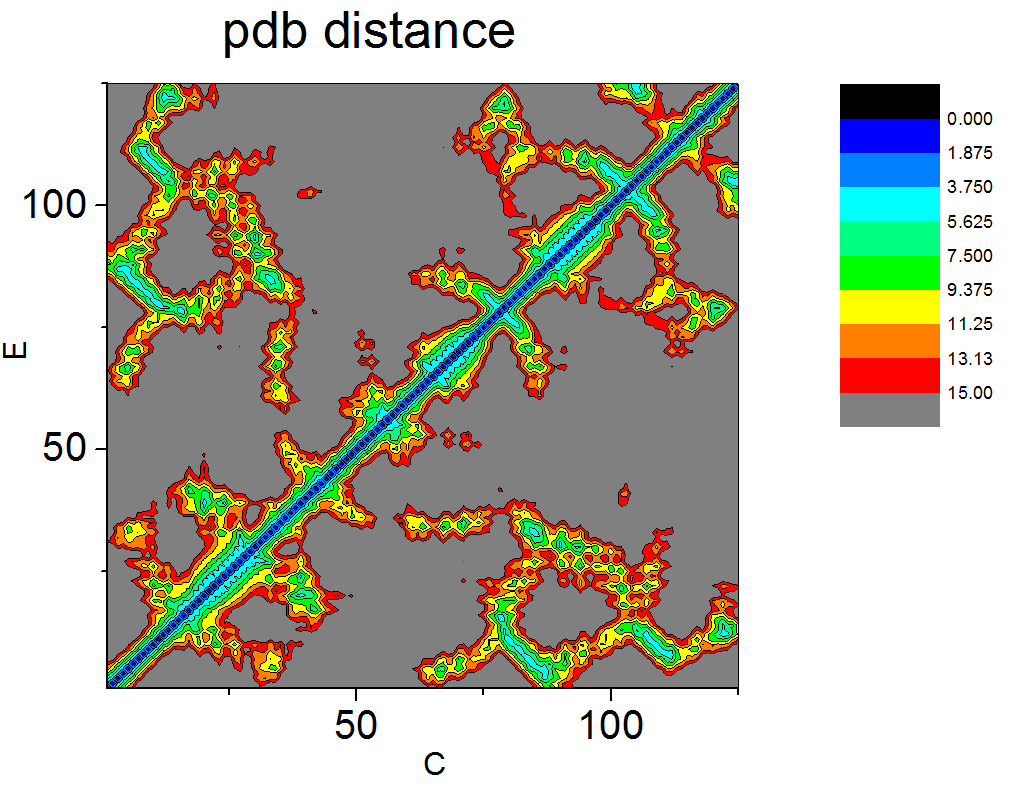
module load biopython

./convert\_alignment.py 5hb7.filt.fas fasta 5hb7.filt.psc

./ccmpred 5hb7.filt.psc 5hb7.filt.mat

to get the contact map of the sequence.

Final results:



For 5HB7, I would say the pattern similarity between the distance and the contact matrix is not as obviously as that for 1nz0D or PCBP1.